



## Heterogeneity in isogenic populations of microorganisms

Pedersen, Anne Egholm

*Publication date:*  
2015

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Pedersen, A. E. (2015). *Heterogeneity in isogenic populations of microorganisms*. Department of Systems Biology, Technical University of Denmark.

---

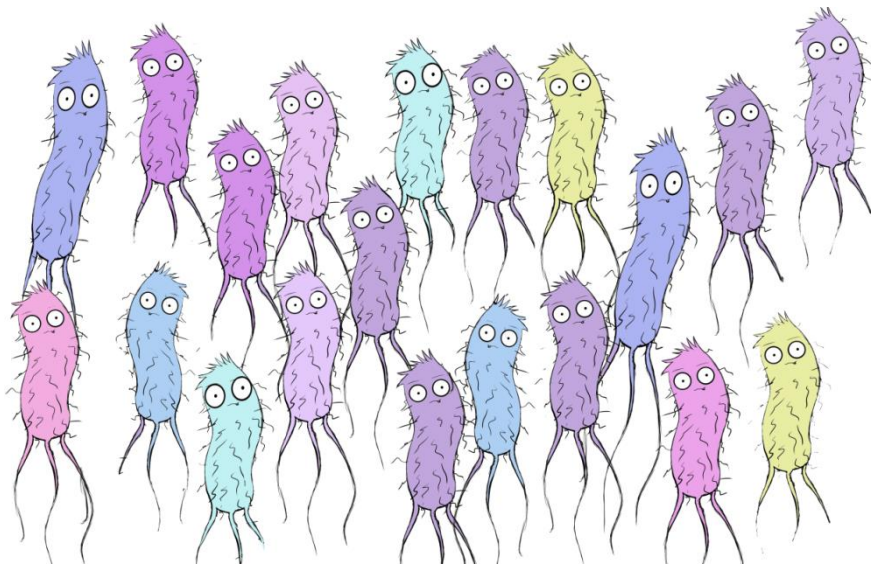
### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Heterogeneity in isogenic populations of microorganisms



**PhD thesis**

**Anne Egholm Pedersen**

Department of Systems Biology  
Technical University of Denmark

# **Heterogeneity in isogenic populations of microorganisms**

## **PhD Thesis**

Anne Egholm Pedersen  
October 2015

Supervisor:  
Assoc. prof. Jan Martinussen

## **Heterogeneity in isogenic populations of microorganisms**

This project was supported by the Danish Council for Strategic Research.

The work of this PhD thesis was carried out at the Department of Systems Biology, Technical University of Denmark. The research was carried out at the Center for System Microbiology (building 301) from July 2009 to July 2010 and from September 2011 to November 2012. From May 2014 to April 2015 the research was carried out in the Metabolic Signaling and Regulation group in building 221. From April 2015 to October 2015 the research was carried out in building 310 in the (new) Metabolic Signaling and Regulation group at the Bacterial Systems Microbiology section. A part of the research was carried out at the Department of Veterinary Biology Disease, University of Copenhagen, from January 2015 to May 2015

Anne Egholm Pedersen

Submitted October 5<sup>th</sup> 2015

Department of Systems Biology  
Technical University of Denmark  
Kgs. Lyngby, Denmark

### **Principal Supervisor:**

Associate professor Jan Martinussen  
Department of Systems Biology  
Technical University of Denmark



## Summary

This work was performed to elucidate heterogeneity in genetically homogenous (isogenic) populations of microorganisms. Phenotypical heterogeneity in populations of isogenic cells are generally accepted, but often ignored and without considering the underlying distribution of the population the mean values for quantifiable variables are used. The reproducibility of an experiment could thus be affected by the presence of subpopulations or high levels of phenotypic variations.

Ole Maaløe and colleagues did in the late 1950'ties observe that the growth rate, RNA, DNA and protein synthesis and cell division changed unsynchronized when a population of *Salmonella typhimurium* was shifted to a different medium supporting another growth rate, hereby setting focus on the importance of balanced growth. In a balanced growing culture every component of the cell increases exponentially by the same constant factor per unit of time. The use of a balanced growing culture is a cornerstone in the Copenhagen School of Bacterial Growth Physiology headed by Ole Maaløe.

Due to the size of the microorganism it is challenging to measure a quantifiable variable in a single cell. However, fluorescence, whether being generated by a fluorescent probe or dye or emitted from a fluorescent protein expressed by the cell, can be detected on a single cell level by microscopy and flow cytometry.

Aiming at quantifying heterogeneity in isogenic populations of microorganisms using flow cytometry fluorescent reporter strains were constructed. A *Bacillus subtilis* reporter strain and a *Lactococcus lactis* reporter strain, both expressing GFP transcriptionally regulated by a strain specific ribosomal promoter, were thus constructed. The reporter strains were validated on a population level and the GFP expression were determined to be growth rate regulated. The growth rate regulated GFP expression could be used as a measure for single cell growth rate and cell-to-cell growth rate variability was investigated by the use of flow cytometry.

Analysis of the *B. subtilis* reporter strain clearly showed that the smallest degree of population heterogeneity was detected when the culture had been propagated according to the guidelines of the Copenhagen School of Bacterial Growth Physiology.

The *L. lactis* GFP reporter strain was more challenging to analyze. The population profile for this reporter strain was shown to be dependent on the type of medium. Chemically defined medium used for propagation of *L. lactis* resulted in a population profile which was problematic to analyze in relation to cell-to-cell growth rate variability.

To investigate population heterogeneity in different types of microorganisms a *Saccharomyces cerevisiae* GFP reporter strain was analyzed. In this strain, a ribosomal protein promoter regulated the GFP expression. It was investigated on a single cell level, whether the use of balanced growth would decrease the population heterogeneity in the organism. The degree of heterogeneity was similar in the samples taken from the balanced growing culture. However, it was observed that balanced growth did not decrease the degree of heterogeneity in a population of *S. cerevisiae*.

Further approaches to investigate population heterogeneity were also investigated. Analysis of GFP expressed from an inducible promoter showed that the degree of heterogeneity was slightly higher at intermediate inducer concentrations. Additionally, the effect of thermal stress on phenotypic heterogeneity was addressed by inflicting a heat stress to the *B. subtilis* GFP reporter strain. The increase in incubation temperature transiently increased population heterogeneity.

Heterogeneity was investigated on a population level using a quantitative analysis of phenotypic variation in methicillin-resistant *Staphylococcus aureus* (MRSA). The results obtained showed that the physiological state of the culture used for the analysis affected the phenotypic variation, as heterogeneity was more pronounced when the analysis was conducted with a stationary phase culture compared to an exponentially growing culture. Moreover, the analysis showed that the effect of cell physiology was much more pronounced in the absence of the generally proposed addition of sodium chloride to the medium.

In summary population heterogeneity in isogenic populations of microorganisms was investigated using different approaches. The primary approach used in this work was single cell analysis by flow cytometry. In the cell-to-cell growth rate variability analysis process it became obvious that procedures for quantification of heterogeneity were lacking. One of the main outcomes of this project was thus the development of procedures for non-statisticians on how to use flow cytometry data to quantify heterogeneity. We are confident that the results presented in this work, will assist the scientific community in quantifying heterogeneity. Finally, the findings in this study signify the importance of a proper experimental setup in order to achieve reproducible and hence valid data.

## Dansk resumé

Formålet med dette projekt var at belyse heterogenitet i populationer af genetisk identiske mikroorganismer. Fænotypisk heterogenitet i populationer af genetisk identiske mikroorganismer er generelt accepteret. Denne heterogenitet er alligevel ofte ignoreret og middelværdier for kvantificerbare variable anvendes, uden at overveje den underliggende fordeling. Tilstedeværelsen af subpopulationer eller stor fænotypisk variation kan derfor påvirke reproducerbarheden af forsøget.

Ole Maaløe og hans kollegaer publicerede i slutningen 1950'erne resultater der viste, at hvis en population af *Salmonella typhimurium* blev flyttet til et nyt vækstmedium, og derved fik en ny væksthastighed, ændredes væksthastigheden, syntese hastigheden af RNA, DNA og protein samt hastigheden hvormed cellerne delte sig - på forskellige tidspunkter efter skiftet. Herved satte denne gruppe forskere fokus vigtigheden af balanceret vækst. I en kultur i balanceret vækst øges hver komponent i cellen eksponentielt med den samme konstante faktor per tidsenhed.

På grund af størrelsen af mikroorganismer er det udfordrende at måle en kvantificerbar variabel i en enkelt celle. Mikroorganismen kan gøres fluorescerende, enten ved brug af en fluorescerende probe eller farve, eller ved at få mikroorganismen til at udtrykke et fluorescerende protein. Denne fluorescens kan detekteres på enkeltcelleniveau ved mikroskopi og flow cytometri.

En *Bacillus subtilis* samt en *Lactococcus lactis* reporterstamme blev fremstillet for at kunne kvantificere heterogenitetsniveauet i populationer af genetisk identiske mikroorganismer. I hver af reporterstammerne blev et grønt fluorescerende protein (GFP) udtrykt. Genet som koder for det fluorescerende protein blev transkriptionelt reguleret af en stammespecifik ribosomal RNA promotor. Reporterstammerne blev valideret på populationsniveau, og det blev påvist at ekspressionen af GFP blev reguleret af væksthastigheden. Den væksthastighedsregulerede GFP ekspression kunne derfor anvendes som en tilnærmelse for den enkelte celledes væksthastighed. Variationen i de enkelte celledes væksthastighed kunne således undersøges ved brug af flow cytometri.

Enkeltcelleniveau-analyse af den fluorescerende *B. subtilis* stamme viste klart, at det laveste heterogenitetsniveau blev opnået når bakteriekulturen voksede balanceret.

Analyse af *L. lactis* GFP reporter stammen var mere udfordrende. Populationsprofilen for denne reporter stamme viste sig at være afhængig af typen af vækstmedium. Kemisk defineret medium resulterede i en population profil som var problematisk at analysere i forhold bestemmelsen af de enkelte celledes væksthastigheder.

For at undersøge heterogenitetsniveauet i forskellige typer af mikroorganismer blev en *Saccharomyces cerevisiae* GFP reporter stamme også analyseret på enkeltcelleniveau. I denne stamme blev GFP reguleret af en promoter tilhørende et ribosomalt protein. Også for denne organisme blev det undersøgt om balanceret vækst vil resultere i et lavere niveau af fænotypisk variation. Resultatet viste at variationen mellem prøverne var mindst, hvis stammen blev voksende balanceret, men at brugen af balanceret vækst ikke resulterede i et lavere heterogenitetsniveau i en *S. cerevisiae* kultur.

For at belyse heterogenitetsniveauet i en population af genetisk identiske mikroorganismer på flere måder blev bl.a. udtrykket af GFP, reguleret af *LacZ* promotoren, målt på enkeltcelleniveau. Analysen viste, at induktion med IPTG koncentrationer som ikke resulterede i en maksimal GFP ekspression forøgede heterogenitetsniveauet. En analyse af hvorledes varmemstress påvirkede heterogenitetsniveauet blev også lavet. En brat, men forbigående, stigning i heterogenitetsniveauet blev observeret for *B. subtilis* reporterstrammen, umiddelbart efter at kulturen blev udsat for varmemstress.

En kvantitativ analyse af fænotypisk variation i methicillin resistente *Staphylococcus aureus* (MRSA) blev foretaget på population niveau. En populationsanalyse-profil blev lavet for at undersøge om variation i stammens resistensniveau blev påvirket af om kulturen, som blev brugt til analysen var eksponentielt voksende eller i stationær fase. Resultaterne viste at brugen af hhv. en eksponentiel voksende kultur og en stationær fase kultur påvirkede variationen i stammens resistensniveau. Resultaterne viste desuden, at hvis analysen blev lavet på et vækstmedium uden tilsat natriumklorid blev forskellen mellem brugen af en eksponentielt voksende kultur i forhold til brugen af en kultur i stationær fase mere udtalt.

Heterogenitet i populationer af genetisk identiske mikroorganismer blev undersøgt ved forskellige fremgangsmetoder. Den primære fremgangsmåde var analyse på enkeltcelleniveau ved brug af flow cytometri. I processen med at undersøge niveauet af variation mellem de enkelte cellers væksthastighed blev det tydeligt at der manglede procedurer til at kvantificere heterogenitetsniveauet. Et af de vigtigste resultater var derfor udviklingen af procedurer til ikke-statistikere til at kvantificere heterogenitet ved brug af flow cytometri data. Resultatet af dette arbejde føler vi os sikre på kan bruges i forskningsmiljøet til at kvantificere heterogenitet. Endeligt viser resultaterne vigtigheden af en korrekt forsøgsopstilling for at kunne opnå reproducerbare og valide resultater.

## Acknowledgement

First of all I would like to thank my supervisor Associate Professor Jan Martinussen for giving me this special opportunity to conduct research within the interesting field of molecular microbiology and for your valuable guidance throughout my entire PhD project.

I would also like to thank Associate Professor emeritus Flemming Hansen and Associate Professor Mogens Kilstrup. Flemming, thank you for always taking the time to look at my genetically modified more or less fluorescent bacteria in the microscope, for introducing me to the spectrofluorometer and for the educative teamwork we had trying to construct *E. coli* GFP reporter strains using recombineering. Mogens, thank you for taking the time to encourage me and being positive regarding my results, for being inspiring and for giving me the opportunity to work with the highly relevant topic - antibiotic resistance in *S. aureus*.

Huge thanks to Maria Amalie Seier-Petersen and Camilla Thyregod. Maria for proofread this thesis and being supportive during the interesting process of the thesis writing. Camilla for the fruitful and educational teamwork we had in connection with the statistical analysis of the flow cytometry data. Also thanks to Kristine Marie Søllingvraa Mikkelsen for improving the Danish summary.

I am grateful for the pleasant company of my many colleagues in the groups I have been a part of during the time of my PhD study, you have all made it a good experience coming to work. Also thanks to Mikkel, Marc and Jacob, “my” three students, it was fun and educational to work together with you. Additionally thanks to Marzanna for always being helpful in the lab.

Finally a special thanks to my mother for always having time to listen and being encouraging towards my aims and dreams. A special thanks to my father for being optimistic and supporting about my work. Last but definitely not least thanks to Morten for being supportive and encouraging, I could not have done this without you. I love you positive attitude and how you can always make me laugh. Also thank you for develop the software “AnnesFileConverter 1” making it possible for me to read the data produced in the spectrofluorometer using Excel and “AnnesFileConverter 2” making it possible to convert data from text files obtained from “Flowing software” to usable csv files.

*To my beautiful children Frey, Siri and Bjørk*

Anne Egholm Pedersen

October, 2015

## Outline of the thesis

The thesis will be divided into nine sections. First a general introduction will be given in **section 1**. This section will include an introduction to phenotypic variation in populations of genetically identical cells, an introduction to the concept of balanced growth and information about the green fluorescent protein (GFP) and consideration using fluorescent proteins for quantitative analysis – concepts and building block which in the following sections will be used in this study of heterogeneity in isogenic populations of microorganisms. In **section 2** will the Gram-positive GFP reporter strains (*B. subtilis* and *L. lactis* reporter strains, respectively) be introduced. The reporter strains were constructed with the purpose to study population heterogeneity on a single cell level by flow cytometry. This section will also include information on flow cytometry data analysis, information about how the flow cytometry data will be displayed and information on the parameter “specific cellular fluorescence”.

**Section 3** includes the manuscript “Heterogeneity at the single cell level in homogenous cultures”. The results presented in this manuscript elucidate heterogeneity in populations of the *B. subtilis* GFP reporter strain. Flow cytometry analysis of the reporter strain was used elucidates how the use of balanced growing cultures affects the degree of population heterogeneity.

The point of departure in each of the following sections (section 4 – 8) will be a specific microorganism and population heterogeneity will, in these sections, be investigated using different approaches. The microorganism in focus in **section 4** is the lactic acid bacteria *L. lactis*. Flow cytometry results obtain by analysis of a green fluorescent reporter strain (introduced in section 2) will be used to elucidate cell-to-cell variability in cultures of *L. lactis*. Wanting to extend the investigation of heterogeneity in isogenic cultures to various microorganisms, single cell analysis of the unicellular eukaryote *S. cerevisiae* was made; the results are presented in **section 5**.

A population of the *B. subtilis* GFP reporter strain (introduced in section 2) was exposed to an abrupt increase in incubation temperature, to investigate how thermal stress affected phenotypic variation, the results is shown in **section 6**. In **section 7** are the results obtained from a preliminary analysis of GFP expression, regulated by the native *lacZ* promoter in *E. coli* presented. The topic of **section 8** is quantitative analysis of phenotypic variation in methicillin-resistant *S. aureus* (MRSA). Results on population analysis profiles will be presented after a more general introduction to MRSA and the quantitative analysis method used. A closing discussion and concluding remarks can be found in **section 9**.

The Appendix in **section 11** includes the manuscript “Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry” (section 11.1). The appendix also includes supplementary materials to section 3, section 4, section 5 and section 6.

# Table of content

<b>Summary .....</b>	<b>I</b>
<b>Dansk resumé .....</b>	<b>III</b>
<b>Acknowledgement.....</b>	<b>V</b>
<b>Outline of the thesis .....</b>	<b>VI</b>
<b>Table of content.....</b>	<b>VII</b>
<b>1 General introduction .....</b>	<b>1</b>
1.1 Phenotypic variation and cellular noise.....	1
1.2 Cell physiology and growth phases .....	3
1.3 The Copenhagen School of Bacterial Growth Physiology .....	4
1.4 Quantitative measurements using green fluorescent protein .....	6
<b>2 Obtaining analyzable fluorescence and flow cytometry analysis.....</b>	<b>8</b>
2.1 Flow cytometry data analysis .....	8
2.2 Specific cellular fluorescence.....	11
2.3 Strain construction .....	12
2.4 Materials and methods strain construction .....	13
2.4.1 Strains, plasmids and primers .....	13
2.4.2 Construction of pAE30 and chromosomally integration in <i>B. subtilis</i> .....	15
2.4.3 Construction of pAE28 and chromosomally integration in <i>L. lactis</i> .....	16
2.5 Results .....	16
<b>3 Heterogeneity at a single cell level in homogenous cultures.....</b>	<b>20</b>
3.1 Abstract .....	20
3.2 Introduction.....	21
3.3 Materials and Methods .....	23
3.4 Results .....	25
3.4.1 Expression of GFP from $P_{mm}$ -gfp fusions correlate with the growth rate at the population level.....	25

3.4.2	Distribution of individual growth rates within balanced cultures .....	27
3.4.3	The level of heterogeneity is determined by the history of the inoculate .....	33
3.4.4	Copenhagen School of Bacterial Growth Physiology revisited .....	37
3.5	Discussion .....	41
3.6	References .....	42
<b>4</b>	<b>Heterogeneity in populations of <i>L. lactis</i>.....</b>	<b>45</b>
4.1	Introduction.....	45
4.2	Materials and methods – <i>L. lactis</i> .....	46
4.2.1	Media .....	46
4.2.2	Experimental setup for making a balanced growing <i>L. lactis</i> culture .....	46
4.2.3	Fluorescence microscopy, spectrofluorophotometry and flow cytometry.....	46
4.3	Results and discussion .....	48
4.3.1	GFP expression in the fluorescent <i>L. lactis</i> reporter strain was growth rate regulated.....	48
4.3.2	Population heterogeneity at different growth rates.....	50
4.3.3	Complex medium changes the shape of the population .....	54
4.3.4	Population profiles complicate the analysis.....	61
<b>5</b>	<b><i>S. cerevisiae</i> single cell analysis.....</b>	<b>63</b>
5.1	Introduction.....	63
5.2	Materials and Methods .....	63
5.2.1	Strains.....	63
5.2.2	Medium .....	64
5.2.3	Experimental setup .....	64
5.2.4	Fluorescence microscopy, spectrofluorophotometry and flow cytometry.....	64
5.3	Results and discussion .....	65
5.3.1	<i>S. cerevisiae</i> GFP reporter strain expressed growth rate regulated fluorescence on a population level .....	65
5.3.2	Large standard deviation in single cell growth rates for the exponentially growing culture ..	66
<b>6</b>	<b>Thermal stress introduces phenotypic variations in <i>B. subtilis</i> .....</b>	<b>72</b>
6.1	Introduction.....	72
6.2	Materials and Methods .....	73
6.2.1	Media .....	73
6.2.2	Fluorescence microscopy, spectrofluorophotometry and flow cytometry.....	73



6.3	Results and Discussion .....	73
6.3.1	Growth proceeded at similar rate after exposure to continuously thermal stress .....	73
6.3.2	Thermal stress increases population heterogeneity .....	75
<b>7</b>	<b>Single analysis of a <i>E. coli</i> strain expressing GFP under regulation of an inducible promoter</b> .....	<b>79</b>
7.1	Introduction.....	79
7.2	Materials and methods – <i>E. coli</i> .....	80
7.2.1	Strains.....	80
7.2.2	Medium .....	80
7.2.3	Spectrofluorophotometry and flow cytometry .....	80
7.3	Results and discussion .....	81
7.3.1	IPTG induction of the <i>lacZ</i> promoter observed on a population level .....	81
7.3.2	The level of population heterogeneity was constant when the <i>LacZ</i> promoter was fully induced .....	82
<b>8</b>	<b>Stationary phase stress and NaCl addition induces homogeneity in resistance profiles from heterogeneous <i>Staphylococcus</i> cultures.</b> .....	<b>87</b>
8.1	Introduction.....	87
8.2	Materials and methods.....	88
8.2.1	Strain .....	88
8.2.2	Media and growth conditions .....	89
8.2.3	Population analysis profiles .....	89
8.2.4	Agar diffusion spot test .....	89
8.3	Results .....	90
8.3.1	Induction of a homogenous profile by stationary phase stress in <i>S. aureus</i> JE2 cultures ....	90
8.3.2	NaCl induces a homogenous profile in balanced JE2 cultures .....	90
8.3.3	The slope of PAP curves reveals underlying populations .....	91
8.4	Discussion .....	92
<b>9</b>	<b>Discussion and concluding remarks</b> .....	<b>94</b>
<b>10</b>	<b>References</b> .....	<b>100</b>
<b>11</b>	<b>Appendix</b> .....	<b>112</b>
11.1	Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry .....	112

11.1.1	Abstract.....	113
11.1.2	Introduction .....	113
11.1.3	Materials and Methods .....	115
11.1.4	Results.....	116
11.1.5	Discussion .....	126
11.1.6	References .....	128
11.1.7	Supplementary material.....	129
11.2	Supplementary materials - section 3.....	139
11.2.1	Heterogenetic determined in populations of <i>B. subtilis</i> growing at three different rates .....	139
11.2.2	Heterogeneity determined in populations of <i>B. subtilis</i> growing in complex or chemically defined medium .....	143
11.2.3	<i>B. subtilis</i> cultivated in glucose chemically defined medium inoculated with either a stationary phase pre-culture or an exponentially growing pre-culture .....	147
11.3	Supplementary materials – section 4.....	155
11.3.1	Heterogenetic determined in populations of <i>L. lactis</i> growing at three different rates .....	155
11.3.2	Heterogeneity determined in populations of <i>L. lactis</i> growing in complex or chemically defined medium .....	159
11.4	Supplementary material – section 5.....	161
11.5	Supplementary material – section 6.....	171

***The difference between “exponential phase” and “balanced growth” is the difference between watching apples fall and thinking of gravity.***

***Moselio Schaechter 2006***

# 1 General introduction

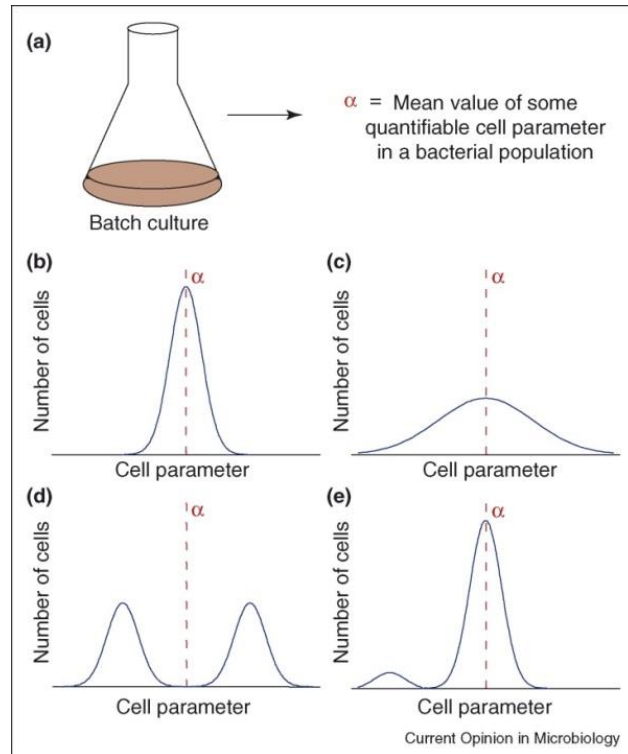
## 1.1 Phenotypic variation and cellular noise

Phenotypic differences within a population of genetically identical bacteria increases survival in the ever-changing microenvironments found in nature. Phenotypic variation is caused by numerous factors, in all probability also some which we have not encountered yet. Stochastic phenotype-switching and phenotypic noise contributes to the cell-to-cell variability present in populations of genetically identical cells. Phenotypic diversity is, from the bacterial populations' point of view, advantageous, as it increases the survival, of at least a part of the population, in case the environment should change. In situations of stress, heat, lack of energy and nutrition or in the presence of  $\beta$ -lactam antibiotics stochastic phenotype-switching mechanisms can occur such as the persistence mechanism in *E. coli* (Balaban *et al.*, 2004), sporulation in *B. subtilis* (Chung *et al.*, 1994; Hoch, 1993), swarming motility (Kearns & Losick, 2005) and natural competence found in various bacteria (Solomon & Grossman, 1996) including *B. subtilis* (reviewed in Dubnau, 1991) generating phenotypic diversity.

Phenotypic noise, generated by stochasticity in gene expression, which generates dispersion of gene product concentrations, entail multiple phenotypes in populations of isogenic bacteria (Eldar & Elowitz, 2010; Elowitz *et al.*, 2002; Levsky *et al.*, 2002; Ozbudak *et al.*, 2002). The stochasticity in gene expression is per definition caused by two types of noises – intrinsic and extrinsic noise. Intrinsic noise is caused by molecular processes, which directly involves production or function of a gene product (promoter activation, mRNA and protein production and decay). Extrinsic noise is caused by fluctuations in the amounts of other cellular components that do not directly involve gene expression, but has the potential to affect gene expression through molecular interactions e.g. numbers of RNA polymerase, quantity of a given repressor protein, the stage in the cell cycle and the cell environment (Elowitz *et al.*, 2002; Swain *et al.*, 2002).

The unavoidable cell-to-cell phenotypic variation in populations of isogenic microorganisms affects the productivity of large scale fermentations. The increase in population heterogeneity caused by feeding zones, pH gradients and oxygen gradients (all extrinsic sources of noise) may explain the lower productivities and biomass yields obtained for cultivation in large scale fermentations, in comparison to well mixed lab-scale fermenters (Bylund *et al.*, 1998; Enfors *et al.*, 2001; Hewitt *et al.*, 2006). In research laboratories cultures are propagated in i.a. shake flasks and lab-scale fermenters. Analysis of samples taken for qPCR, proteomics and genome sequencing etc. give results which are based on population averages. Results based on analysis at a single cell level have been published: Real-time RNA profiling by measuring promoter activity in a single cell (Le *et al.*, 2005), genome sequencing from single cells (Yilmaz & Singh, 2012; Zhang *et al.*, 2006a) as well as several studies using fluorescent proteins to study phenotypic noise (Elowitz *et al.*, 2002; Golding *et al.*, 2005; Ozbudak *et al.*, 2002). A main theme in the methods papers concerning genome sequencing and real-time RNA profiling from a single cell is how to deal with the small amount of starting

material. For simplicity population averages is used instead of single cell observations, hereby making the assumption that the population of isogenic cells is homogenous with regard to phenotypic variation. The established cell-to-cell variability is thus ignored. Making this assumption makes it possible to use the population average without taken into consideration that the averages does not reflect the underlying distribution, and hence says nothing about the standard deviation for the parameter determined and whether the culture is divided into subpopulations, as illustrated in Figure 1.



**Figure 1** This illustration is borrowed from an article by Dhar and McKinney, 2007. In the figure caption the authors describe how the method of analysis can mask for the underlying heterogeneity within a homogenous bacterial population. Cite from the paper by Dhar and McKinney: "(a) Quantification of a cellular parameter in a batch culture yields a mean value ( $\alpha$ ). This value could reflect very different underlying population distributions, including: (b) Gaussian with a narrow distribution around the mean, (c) Gaussian with a broad distribution around the mean, (d) multimodal distribution comprising subpopulations of equal size, or (e) multimodal distribution comprising populations of unequal sizes". (Dhar & McKinney, 2007)

## 1.2 Cell physiology and growth phases

Ninety seven years ago Burhanan published a paper entitled “Life phases in a bacterial culture” (Buchanan, 1918). In this paper the different phases of a bacterial population life cycle was described; initial stationary phase, positive growth acceleration phase (lag phase), logarithmic growth phase, phase of negatively growth acceleration, maximum stationary phase, phase of accelerated death and logarithmic death phase. Information's which today is part of every microbiology text book (Cooper, 1993). The characteristics of the cells in the different phases have been thoroughly investigated during the many years of research conducted since the publication of this paper. Some of the cell characteristics in lag, exponential and stationary phases will be outlined here.

The adjustments of the cells in the different growth phases affect the changes occurring in the next phase; lag phase cells are a product of stationary phase cells adjusting to a new environment, increasing biomass and preparing cell division (Madar *et al.*, 2013). The adjustments of a cell when it has entered stationary phase encounter morphological and physiological changes; the cell volume decreases and the cell shape changes (Kolter *et al.*, 1993; Neidhardt *et al.*, 1990). In *E. coli* alteration in cell wall composition takes place (Cronan, 1968) and translation activity decreases (Wada *et al.*, 1990). The pattern of gene expression changes (Ishihama, 1997, 1999) and osmotic damages accumulates, impacting on both DNA and proteins (Nyström, 2003; Saint-Ruf *et al.*, 2008). Stationary phase cells, in which all these modifications have taken place are more resistant to osmotic stress and heat stress (Givskov *et al.*, 1994; Jenkins *et al.*, 1990). However, once inoculated in fresh medium the gene expression pattern will again be modified as the cells adjusts to a new environment. During the lag phase will the ribosome content in the cells gradually increase, to the level observed in the exponential phase (Rolfe *et al.*, 2012). During the lag phase the biomass will increase and when the cells start to divide the exponential growth phase will be initiated. In the exponential phase will the cells divide at a constant rate depending upon the composition of the growth medium and the incubation temperature. The cell division rate (the growth rate) will in a given medium impose the biochemical composition of the cells and thus the ribosome and RNA content of the cells in that specific medium (Schaechter *et al.*, 1958). Here it is important to note that the rate of rRNA or ribosome synthesis does not determine the growth rate. Rather, the rate of ribosome synthesis responds to the growth rate (Gourse *et al.*, 1996).

The ribosomal promoters regulate the transcription of rRNA operons and proportionality between growth rate and transcriptional activity of the rRNA promoters is expected (Kjeldgaard *et al.*, 1958; Neidhardt & Magasanik, 1960; Schaechter *et al.*, 1958). As ribosome synthesis is energetically costly the transcriptional activity of the ribosomal promoters will comply with the changing translational requirements of the cell, e.g. under different growth rates (Krásný and Gourse 2004).

During exponential growth the cells will have reached the maximum cell size given for that specific growth condition and fluctuation in cell size will, in this growth phase, primarily be caused by cell division. However, observations on single cell level have already decades ago disclosed variation in cell size in homogenous populations (Schaechter *et al.*, 1962; Sharpe *et al.*, 1998). As the medium at some point will be depleted for

nutrients, or it becomes so acidic that further growth is precluded, cell divisions will reaccelerate and the culture will re-enter stationary phase.

### 1.3 The Copenhagen School of Bacterial Growth Physiology

A bacterial culture propagated as a batch fermentation will per definition enter the different growth phases and the phenotype of the cell will be dictated by these. The progression through the bacterial culture life phases was considered inevitable for decades. According to the interesting introduction to the “Schaechter-Maaloe-Kjeldgaard experiments” authored by Stephen Cooper was the concept of steady state growth introduced by Monod in the late 1940'ties. The concept of balanced growth, defined by a condition of steady-state growth with all cellular components increasing at the same rate, was introduced in 1957 by Allan Campbell (Cooper, 1993).

In the late 1950'ties two papers discussing the subject of balanced growth were published. The first paper “Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*” (Schaechter *et al.*, 1958), describes balanced growth at different growth rates. The study presented in this paper emphasizes that no matter how the medium is altered, the average composition of the cells is set by the growth rate. The second paper “The transition between different physiological states during balanced growth of *Salmonella typhimurium*”, by the same authors, was also published in 1958. Results presented in this paper illustrate how a balanced growing bacterial culture, shifted from a medium supporting slow growth to one that leads to a higher rate or vice versa (in either shift-up or shift-down experiments) changes the growth rate, the RNA, DNA and protein synthesis and cell division at non-synchronous after the shift (Kjeldgaard *et al.*, 1958). Kjeldgaard *et al.* describe a shift from minimal salt medium to broth for an exponential growing culture and they observe an abrupt change in biomass, indicating that a new exponential rate was established instantly. The abrupt increase in biomass was shown not to be caused by an increase in cell number (based on colony count), but consequently an increase in cell size. The rates of DNA synthesis and cell division changed more than 20 minutes later (Kjeldgaard *et al.*, 1958). In the preceding paper by Schaechter *et al.* showed that *S. typhimurium* have a given size at a given growth rate (Schaechter *et al.*, 1958). Thus it was observed that a balanced growing culture moved from one state to another, needs a period of time to adjust to the new conditions (Kjeldgaard *et al.*, 1958).

According to the Copenhagen School of Bacterial Growth Physiology biochemical data are only meaningful if the experiment performed to generate the data is made with a population of isogenic cells, growing balanced (Ingraham *et al.*, 1983). Kjeldgaard *et al.* showed in the shift-up experiment, where a balanced minimal salt medium culture was shifted to broth, that the culture became unbalanced (Kjeldgaard *et al.*, 1958). The shift caused the cell division rate and the rates at which biomass increases and RNA and DNA was synthesized to change, but in an unsynchronized manner, hence the culture was not in balance after the shift and needed time to adapt to the new conditions (before a now balanced state was established). When a stationary phase culture is inoculated in fresh medium it can be viewed as a shift up and the unsynchronized state after the

shift could potentially last longer than when an exponentially growing culture experience a medium shift, as the biochemical composition and the gene expression pattern is redefined during the lag phase.

Those who have been educated in the “Copenhagen School” know that the use of a balanced growing culture, a culture which has been growing exponentially for several, are an important basis in every growth experiment. A balanced growing culture should, in a batch cultivation, be growing unrestricted. Unrestricted growth is defined by growth in a medium where the growth rate is limited by the type of nutrients and not by their concentration, as would be the case in a continuous cultivation (Schaechter *et al.*, 1958).

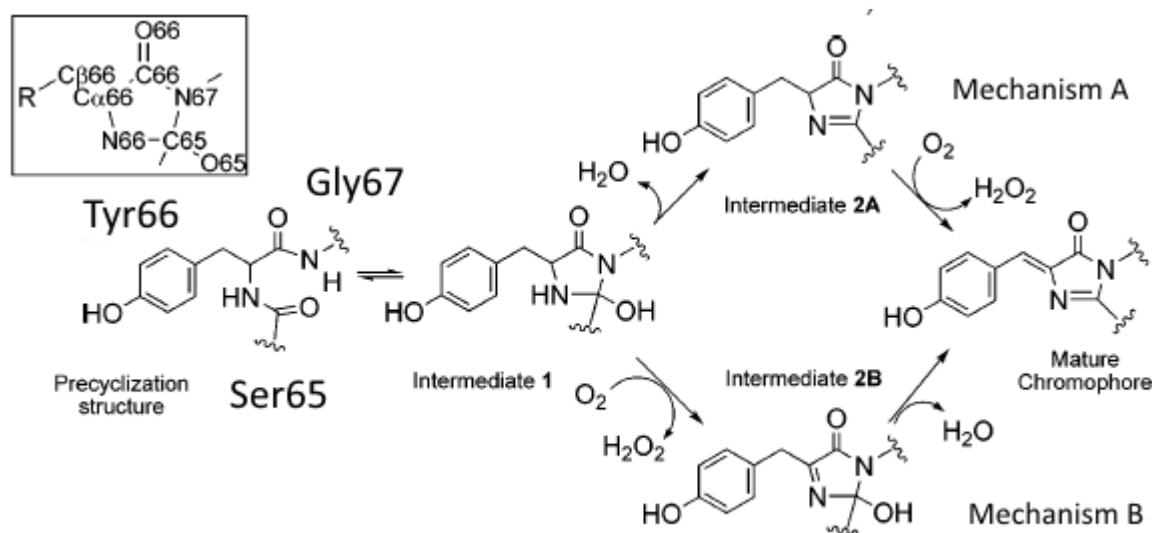
Liquid cultures in which the population are balanced growing must have a low bacterial concentration, so the cells can continue to grow for a long time in a virtually unchanging environment (Schaechter *et al.*, 1958). Having a low bacterial concentration in a liquid culture is the ideal condition of balanced growth and can per definition be prolonged by dilution of the cells into fresh pre-warmed medium (Schaechter *et al.*, 1958). It is important to note that a low bacterial concentration is required to maintain balanced growth. If the medium composition is altered by depletion of a medium component (e.g. an amino acid) the growth rate would be affected and the culture would per definition no longer be balanced growing.

Even though a strict definition of balanced growth require a minimal growth medium where all components are in excess, complex medium was used in the papers from 1958 (Kjeldgaard *et al.*, 1958; Schaechter *et al.*, 1958) to support balanced growth. This could only be attained if the cell density was kept low. In a paper from 2007 Sezonov *et al.* measured the average cell mass (cell concentration per OD<sub>600</sub>) of *E. coli* cells propagated in LB and concluded that LB only supports balanced growth for *E. coli* to an OD<sub>600</sub> of 0.3 (Sezonov *et al.*, 2007). They concluded that changes in growth rate (a specific growth rate supports a specific cell size (Pierucci, 1978; Schaechter *et al.*, 1958)) were caused by depletion of amino acids, as the growth rate slows down gradually after OD<sub>600</sub> of 0.3, reflecting a gradual impoverishment of the medium. Sequential catabolism of amino acids has been determined in tryptone broth for *E. coli* (PrüB *et al.*, 1994). The sequential removal of amino acids causes a decrease in average cell mass (growth rate) and may trigger the stringent response. Even when a single amino acid is missing from the growth medium, uncharged tRNAs are formed and trigger ppGpp synthesis by RelA (stringent response). This induces a response resulting in downregulation of the whole protein transcriptional/translational machineries that apply limits on growth (Lazzarini *et al.*, 1971; Traxler *et al.*, 2008).

The optical density at which an amino acid of a complex medium becomes depleted will depend on the state of the inoculum and how much the inoculum is diluted. The composition of the complex media could potentially be batch dependent and component concentrations could consequently vary from batch to batch. Using a chemically defined medium for supporting balanced growth a specific component could also be depleted if the bacteria concentration exceeded a certain level. In order to obtain a balanced growing culture in a batch fermentation the inoculum must be exponentially growing once inoculated and balanced growth will only continue until the first medium component becomes depleted.

## 1.4 Quantitative measurements using green fluorescent protein

In the Pacific Ocean at the west coast of North America lives the jellyfish *Aequorea victoria*. This jellyfish activates the photoprotein aequorin by releasing calcium ( $\text{Ca}^{2+}$ ). The activated aequorin produces blue light which is transduced to a green fluorescent protein (GFP). GFP emits green light and the jellyfish becomes bioluminescence (Heim & Tsien, 1996; Prasher, 1995). In the center of the tertiary structure of this small 238 amino acid protein is the chromophore encapsulated. The chromophore is formed in an autocatalytic cyclization of the tripeptide Ser65-Tyr66-Gly67. In the presence of oxygen, the chromophore emits green light when excited with blue light (Heim *et al.*, 1994). Two mechanisms have been proposed for the chromophore maturation: A) a cyclization-dehydration-oxidation mechanism and B) a cyclization-oxidation-dehydration mechanism. In Figure 2 the maturation processes is illustrated, according to Timothy Craggs results published by Zhang *et al.* in 2006, favors mechanism B shown in Figure 2, but adds that both mechanisms can run in parallel and that the relative flux through each of the mechanisms pathway may depend on oxygen concentration and the particular FP variant (Craggs, 2009; Zhang *et al.*, 2006b).



**Figure 2** Folding and maturation of GFP happens in three steps. Nucleophilic attack of the amide of Gly67 on the carbonyl of residue 65 (see framed illustration) followed by dehydration and then oxidation, with concomitant production of a hydrogen peroxide (mechanism A). The pathway where intermediate 1 first becomes oxidized (intermediate 2B) and then dehydrated and the chromophore acquires visible fluorescence after one of the  $\text{C}\beta_{66}$  hydrogens is lost (Craggs, 2009). Illustration is borrowed from (Zhang *et al.*, 2006b) and has been modified with information from (Craggs, 2009).

The folding of the protein and the maturation of the chromophore has been estimated to occur with a time constant of 2-4 hours for wild type GFP (Heim *et al.*, 1994). Others have estimated the maturation to take less than an hour (Iizuka *et al.*, 2011). The maturation time of GFPmut1 is estimated to be approximately eight



minutes in *B. subtilis* (Kesel *et al.*, 2013). The maturation time for GFPmut3 is measured to be 450 seconds *in vitro* (Iizuka *et al.*, 2011). However the result was obtained *in vitro* and does not necessarily agree with actual maturation rates *in vivo*. Hebisch *et al.* measured the maturation time of the same GFP variant (GFPmut3) in different *E. coli* strains, and observed that maturation time (determined to be between 4 and 6 minutes) is also strain dependent, but concludes that the differences observed between the strains are within measurements uncertainties (Hebisch *et al.*, 2013). In conclusion the immature proteins are present in the cells minutes before the chromophore matures and fluorescence can be detected. However, based on literature studies it would be difficult to determine the exact maturation time as it depends on GFP variant, experimental condition and bacteria strain.

One of the main problems with wild type GFP is that it has two excitation peaks, a major one at 395 nm and a smaller one at 475 nm. By modification of the tripeptide located in the chromophore GFP variants with optimized abilities to be analyzed by flow cytometry (excitation maximum closer to 488 nm, which is the wavelength most common for the primary flow cytometers lasers) have been created. One of these is GFPmut1 in which the serine at position 66 is substituted by threonine (Cormack *et al.*, 1996). Other changes in the amino acid sequence have resulted in faster folding GFP, yellow fluorescent protein (YFP), cyan fluorescent protein (CFP) and blue fluorescent protein (BFP) (Cubitt *et al.*, 1995; Miyawaki *et al.*, 1999; Rizzo *et al.*, 2004).

While classical approaches based on enzymatic quantification of a reporter (e.g.  $\beta$ -galactosidase) only provide information concerning the gene expression process within the population, fluorescent reporter assays offer high resolution information at the single cell level (Elowitz *et al.*, 2002).

On a population level several studies have shown that measurement of fluorescence intensity is as reliable as a  $\beta$ -galactosidase assay or a chloramphenicol acetyltransferase (CAT) assay. Scholz *et al.* demonstrate that GFP and  $\beta$ -galactosidase expression are equally reliable indicators of *tetA* promoter activity in *E. coli* (Scholz *et al.*, 2000). Albano *et al.* found that GFP and CAT provide equivalent measures of *araBAD* promoter activity in *E. coli* following arabinose-induced expression of a GFP-CAT fusion construct (Albano *et al.*, 1998).

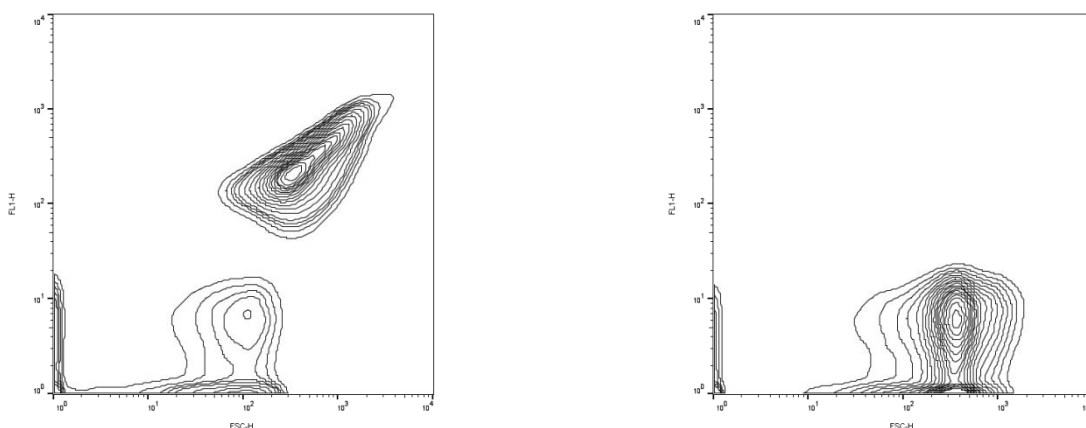
## 2 Obtaining analyzable fluorescence and flow cytometry analysis

To elucidate heterogeneity in homogenous cultures custom designed *B. subtilis* and *L. lactis* reporter strains were constructed. Included in this section is information about the strain construction and information on how the flow cytometry data will be presented throughout this thesis. Descriptive statistics and the parameter specific cellular fluorescent will also be introduced in this section.

Flow cytometry is a powerful tool for single cell analysis and possesses a capacity to analyze many cells per second. Cell size and fluorescence, whether being fluorescence staining or fluorescence produced by expression of a fluorescent protein, can be determined on a single cell level by flow cytometry analysis. Forward scatter (collected co-linearly with the illumination source) is assumed to be proportional to cell size (Christensen *et al.*, 1995; Heddal *et al.*, 1994). Since it generally accepted that forward scatter is proportional to cell size, even though the scatter signal potentially also is influenced by cell structure and chemically composition (Julià *et al.*, 2000; Robertson *et al.*, 1998). Forward scatter will be used to define cell size throughout this work.

### 2.1 Flow cytometry data analysis

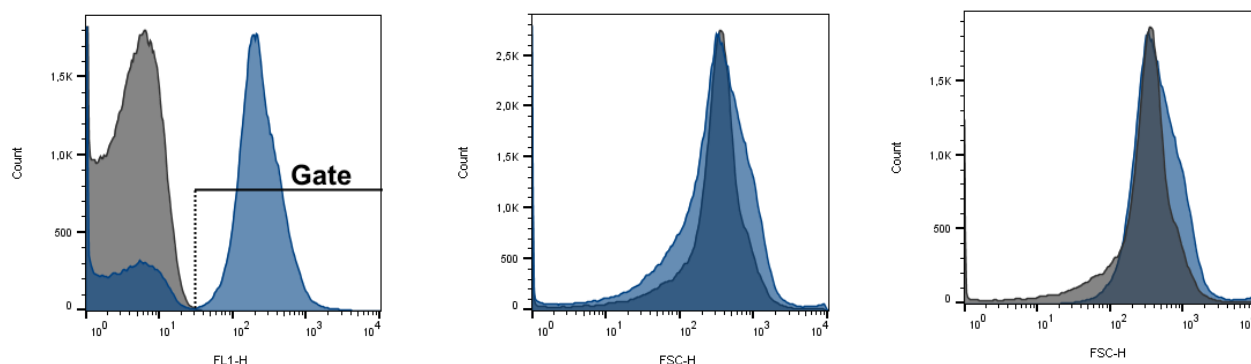
Ungated flow cytometry data will in the following sections be displayed in contour plots or as histograms. The contour plots used have 5 % probability contours, meaning that 5 % of the cells lie between each contour. The contour plots are illustrative and provide an intuitive impression of the population profile and the degree of heterogeneity. Two contour plots can be seen in Figure 3; a green fluorescent *L. lactis* reporter strain (left) and the associated non-fluorescent control strain (right). A clear difference between the two strain was observed (the flow cytometry settings used were identical for the analysis of the two *L. lactis* strains). The horizontal axis represents forward scatter per cell (FSC-H) whereas the vertical axis represents fluorescence per cell (FL1-H). Both axes are in a four decade logarithmic scale.



**Figure 3** Contour plot with 5 % probability contours. Forward scatter per cell (FSC-H) is represented by the horizontal axis. Fluorescence per cell (FL1-H) is represented by the vertical axis. Both axes are in a four decade logarithmic scale. The samples presented here show population profiles for the *L. lactis* strains when propagated in complex medium. The population profile of the green fluorescent reporter strain (AE072) is shown in the plot to the left. The population profile for the control strain (AE071) is shown in the plot to the right. The flow cytometry settings were identical for analysis of the two strains, the fluorescence signal was amplified to an extent where the cells of the control strain clearly possessed detectable fluorescence, even though they were not expressing GFP.

Overlay histograms (or similar, e.g. offset histograms) can provide a quantitative view of the variation for one parameter and provide information of skewness and whether subpopulations are present. The variation within a dataset can also be displayed in a box-plot. Box-plots offer an intelligible quantitative way to compare large amounts of data. With a view to compare population heterogeneity within samples, between samples and between different experimental setups box-plots are an ideal choice. The box-plot shown in the following section has a 50 % interval placed in the “box” and the ends of the whiskers represent 2.5 percentile and 97.5 percentile, respectively. Above the box-plots are growth curves shown, each diamond symbol in the growth curves can be linked to a box-plot placed in the same sub-section of the figure.

Analysis of control strains were used to isolate cells of the reporter strains containing GFP. Cells expressing GFP, but having damaged/permeabilized membranes, leaking of GFP was observed using microscopy (data not shown). Histogram gating was used to exclude cells with permeabilized membranes and cell debris. All data shown as box-plot and used for calculating descriptive statistics are gated events. The events were gated based on their fluorescence. To illustrate how this was done an overlay histogram showing the control strain and the reporter strain is displayed in Figure 4(left).



**Figure 4** Overlay histogram showing the distribution of fluorescence per cell [FL1-H] (left) and forward scatter per cell [FSC-H] (middle and right) for the *L. lactis* GFP reporter strain (blue histograms) and the control strain (grey histograms). The data is the same as presented in Figure 3. The peak height was normalized by the software (FlowJo). The black bar in the overlay histogram to the left represents the gate used to isolate the events of interest for further analysis (box-plot and descriptive statistics). The overlay histogram in the middle shows the forward scatter for the ungated data for the GFP reporter strain. The overlay histogram to the right shows the gated data for the GFP reporter strain. It is the ungated data for the control strain which is shown in the three histograms.

Flowing Software (v2.5.1, Turku Centre for Biotechnology, Turku, Finland) software was used for gating, before statistical analysis was done in JMP® (statistical software from SAS).

Descriptive statistics were calculated based on the gated data. The gated data collected for a sample will be presented in a box-plot and the descriptive statistics; mean and median value, the standard deviation, the coefficient of variation and the interquartile range, calculated using JMP®, will be shown in scatter plots. Above the box-plots is growth curves shown. Each scatter point can be linked to a diamond symbol in a growth curve above the box-plots. Each box-plot can likewise be linked to a diamond symbol. E.g. can the descriptive statistics in Figure 13A be linked to the growth curves above the box-plots in Figure 11A, the references will be given in the figure captions. The contour plots can likewise be linked to diamond symbols in the growth curves shown above the box-plots.

## 2.2 Specific cellular fluorescence

Investigating population heterogeneity at a single cell level using GFP, the nature of GFP, cell physiology and the combination of these, should be included in the interpretation of the results. The GFP, unless provided with a peptide tail which makes it a target for intracellular tail-specific proteases (Andersen *et al.*, 1998), is a stable protein, which would stay fluorescent within the time of the experiments. The half-life of a stable GFP has been estimated to be more than 24 hours (Andersen *et al.*, 1998).

Custom designed reporter strains, in which the expression of GFP was transcriptionally regulated by a ribosomal promoter, were constructed with the objective to study phenotypic variation in populations of isogenic cells. The number of ribosomes in the cells is set to match the potential for protein synthesis in the culture medium used (Molin & Givskov, 1999; Maaløe & Kjeldgaard, 1966). Using an ribosomal promoter to transcriptionally regulate the GFP expression in the reporter strains made it possible to use GFP fluorescence as a measure of the ribosomal promoter activity in the cell during exponential growth.

The concentration of matured GFP in the cell, and not the total fluorescence of the cells, was used as a quantitative measure of single cell growth rate. The concentration of matured GFP was calculated by dividing the fluorescence measured for a single cell by the cell size, determined by forward scatter for that particular cell. The concentration of matured GFP will in this work be referred to as specific cellular fluorescence. Taken the cell size into consideration, when using the cellular fluorescent to reflect the activity of the ribosomal promoter, was important for several reasons. In *B. subtilis* faster growing cells are larger (Sargent, 1975; Sharpe *et al.*, 1998), so if cells at different growth rate should be compared the cell size would have an effect on the GFP content of the cells. The activity of the ribosomal promoters in cells growing exponentially at the same rate is similar in an elongated cell which is preparing for division and in a smaller cell just gone through a cell division. In both cases they will be detected as a single event by the flow cytometer. By using the GFP concentration, instead of the total cellular GFP, this will be taken into consideration.

The reporter strains were constructed to observe population heterogeneity during exponential growth. Results obtain from analysis of samples taken during lag phase or in stationary phase are interesting, but must be interpreted with care as the relation between cell division, GFP expression and GFP maturation is not in a quasi-steady state as would be the case during exponential growth. Cells in stationary phase were fluorescent as GFP did accumulate in the cell during exponential growth. However, the intensity of the fluorescence could not, in stationary phase, be linked to the activity of the ribosomal promoter regulating the expression of GFP.

The distribution of fluorescence per cell and forward scatter per cell, the underlying parameters to specific cellular fluorescence, can be visualized in box-plots in supplementary materials, together with the associated descriptive statistics.

## 2.3 Strain construction

Wanting to investigate population heterogeneity in cultures of isogenic cells using flow cytometry various parameters could be used to elucidate this; heterogeneity in cell size under different growing conditions, heterogeneity in production of a fluorescent fusion protein and heterogeneity in antibiotic resistance. One of the objectives of this study was to investigate heterogeneity in single cell growth rate by determination of expression of a fluorescent protein; hence reporter strains expressing a fluorescent protein under regulation of a ribosomal promoter were constructed. GFP was an obvious choice to use as reporter protein; it has been used for many applications and in various organisms and the GFP chromophore is excited with light at 488 nm. 488 nm light can be generated by a blue laser; the primary laser type in many flow cytometers.

Several demands were outline for the design of the reporter strains: GFP expression should be regulated by a ribosomal promoter, *gfp* should be present in a single chromosomally copy and the reporter strains should be fluorescent enough for single cell analysis by flow cytometry.

The initial objective was to analyze population heterogeneity in Gram-positive bacteria, and *B. subtilis* and *L. lactis* were likely candidates, being model organisms in Gram-positive and lactic acid bacteria research. Reporter strains meeting the customized demands were constructed. During the construction of the preliminary strains it became evident that the codon usages and the variant of GFP clearly had an effect on the fluorescence intensity of the cells.

## 2.4 Materials and methods strain construction

### 2.4.1 Strains, plasmids and primers

**Table 1 Strains**

Strain	Relevant characteristic	Reference
<i>E. coli</i> DH5 $\alpha$	F80 <i>lacZ</i> DM15 D( <i>lacZ</i> YA- <i>argF</i> )U169 <i>recA1 endA1 hsdR17 supE44 thi-1 gyrA96 relA1</i>	Laboratory strain
<i>B. subtilis</i> strain 168	<i>trpC2</i>	(Burkholder & Giles, 1947; Spizizen, 1958)
AE003	<i>B. subtilis</i> strain 168 amyE::pAE04	This study
AE044	<i>B. subtilis</i> strain 168 amyE::pDG268neo	This study
AE099	<i>B. subtilis</i> strain 168 amyE::pAE30	This study
<i>L. lactis</i> MG1363	Plasmid-free derivate of strain NCDO 712	(Gasson, 1983)
AE071	MG1363 transformed with pMK1167, empty vector, cured for pLB95	This study
AE064	MG1363 transformed with pAE22, <i>laGFP</i> and <i>rrnC</i> promoter, cured for pLB95	This study
AE072	MG1363 transformed with pAE28, <i>gfp-opt</i> and <i>rrnC</i> promoter, cured for pLB95	This study

**Table 2 Plasmids**

Plasmid	Relevant characteristic	Reference
pDG268neo	Ap <sup>r</sup> Nm <sup>r</sup> ; a vector for constructing transcriptional <i>lacZ</i> fusions designed to integrate in the chromosomal <i>amyE</i> gene which contains a promoterless <i>lacZ</i> gene and neo resistance cassette in opposite directions flanked by the truncated <i>amyE</i> gene	C. W. Price, Department of Food Science and Technology, University of California, Davis, California, USA.
pAE04	<i>gfp2191</i> from (Hansen & Atlung, 2011) in BamHI:PciI. P <sub>rrnA</sub> in EcoRI:BamHI	This study
pAE30	<i>gfp-opt</i> from (Sastalla <i>et al.</i> , 2009) in BamHI:PciI. P <sub>rrnA</sub> in EcoRI:BamHI	This study
pSW4-GFPopt	Vector containing <i>gfp-opt</i>	(Sastalla <i>et al.</i> , 2009)
pLB86	<i>attP</i> , <i>lacLM</i> , <i>bla</i> , <i>erm</i> - promoterless <i>lacLM</i> integration vector	(Brøndsted & Hammer, 1999)
pMK1167	<i>attP</i> , <i>bla</i> , <i>erm</i> , <i>lacLM</i> – promoterless <i>lacLM</i> integration vector (pLB86 ) modified with a terminator site placed upstream <i>lacLM</i>	(Mogens Kistrup, not published)
pLB95	<i>orf1 tet ori<sub>ts</sub></i>	(Brøndsted & Hammer, 1999)
pAE09	pMK1167 with <i>rrnC(b)</i> , <i>gfp-2191</i> and <i>fbaA</i> leader	This study
pAE22	pAE09 with <i>la-gfp/fbaA</i> leader	This study
pAE28	pAE09 with <i>gfp-opt/fbaA</i> leader	This study



**Table 3 Primers**

Primers	Sequence (5'-3'), restriction sites are underlined	Use
rrnA_fwd	GCTAG <u>AATTC</u> GTTACAACAGCTATCAGCGG	P <sub>rrmA</sub> for pAE30 and pAE04
rrnA_rev	GCTAGGAT <u>CCCC</u> AGTCTTACAGGCAGGTT	P <sub>rrmA</sub> for pAE30 and pAE04
RBS_gfp_fwd_CJ	CG <u>GGATCC</u> AAAGGAGGAAAACATATGTCTAAAGGTGAAGAACTG	<i>gfp2191</i> for pAE04
gfp_rev_CJ	CCATACATGTTTATTTATACAGCTCATGCATGC	<i>gfp2191</i> for pAE04
Fba(BamHI)2_fwd	GCTACT <u>GGATCC</u> GCTAGAGAACAATCAGGCTACATAAGGAGGACATTCTCATGTCAAAAGGTGAAGAATT	<i>gfp-opt</i> for pAE30
laGFP(Pcil)_rev	CGAAGTACATGTTTATTTATATAATTCATCC	<i>gfp-opt</i> for pAE30
rrnC(b)_fwd	CGACGT <u>GTCGACT</u> GACAGTCTTAGAAAC	P <sub>rrnC</sub> for pAE22 and pAE28
rrnC(b)_rev	CGACGTCTGCAGGTATTTAAGTCACCGAAGTG	P <sub>rrnC</sub> for pAE22 and pAE28
Fba(PstI)_fwd	GCTACTCTGCAGAGGTAGGTAAAAAATATTCGGAGGAATTTTGAAATGTCAAAAGGTGAAGAATT	<i>gfp-opt</i> and <i>laGFP</i>
laGFP(HindIII)_rev	CGAAGTAAGCTTTTATTTATATAATTCATCC	<i>gfp-opt</i> and <i>laGFP</i>

*E. coli* strain DH5 $\alpha$  was used for genetic manipulations. The *E. coli* cells were transformed by electroporation as previously described (Sambrook and Russell, 2001). *E. coli* were propagated in lysogeny broth (LB) (Bertani, 1951), and selected on LB plates supplied with 100  $\mu$ g/ml ampicillin. Transformation of *B. subtilis* was performed as described by (Boylan *et al.*, 1972), transformants were selected on LB plates supplemented with 5  $\mu$ g/mL neomycin. *L. lactis* cells were transformed by electroporation as previously described (Holo & Nes, 1989). *L. lactis* transformants was selected on M17 (Terzaghi & Sandine, 1975) plates supplied with 1 % glucose, and 5  $\mu$ g/ml erythromycin.

Restriction was carried out using Fermentas restriction enzymes and buffer system. Ligation was performed with T4 DNA ligase (New England Biolabs®Inc.) Gel electrophoresis and transformation of competent *E. coli* were carried out as described by (Sambrook & Russell, 2001).

#### 2.4.2 Construction of pAE30 and chromosomally integration in *B. subtilis*

*gfp-opt* was amplified from pSW4-GFPopt using the primers Fba(BamHI)2\_fwd and laGFP(Pcil)\_rev. A 40 bp fragment of the FbpA leader, including the strong ribosomal binding site present in the leader, was placed upstream *gfp-opt*. The resulting PCR fragment was digested with BamHI and Pcil and ligated into the *B. subtilis* integrative vector pDG268neo, digested with the same restriction enzymes. The recombinant plasmid was digested with EcoRI and BamHI and ligated together with a PCR fragment made by amplifying

the  $P_{rrnA}$  promoter (16S rRNA promoter, including both P1 and P2 and 257 bp downstream +1) using the primers *rrnA\_fwd* and *rrnA\_rev*. The recombinant plasmid (pAE30) was amplified in *E. coli* DH5 $\alpha$ . Plasmid pAE30 was integrated into the *B. subtilis* 168 *amyE* gene by a double cross-over recombination through the *amyE* N-terminal and *amyE* C-terminal parts located on the pDG268neo vector. The plasmid was linearized by digestion with XhoI prior to transformation in order to ensure chromosomal integration by a double cross-over event. The neomycin-resistant strain AE099 was obtained. The control strain AE044 was obtained by transforming *B. subtilis* strain 168 with XhoI-linearized pDG268neo.

### 2.4.3 Construction of pAE28 and chromosomally integration in *L. lactis*

The ribosomal promoter  $P_{rrnC}$  was amplified from the chromosome of *L. lactis* subsp. *cremoris* MG1363 using the primers *rrnC(b)\_fwd* and *rrnC(b)\_rev*. 146 bp downstream of +1 and 241 bp upstream the 16S rRNA gene was amplified. The PCR fragment (*rrnC(b)\_fwd/rrnC(b)\_rev*) was integrated in Sall and PstI restriction sites of plasmid pMK1167. *gfp-opt*, a gene encoding a GFP variant (GFPmut1) codon optimized for *Bacillus anthracis* (Sastalla *et al.*, 2009) was amplified from pSW4-GFPopt using the primers *Fba(PstI)\_fwd* and *laGFP(HindIII)\_rev*, changing the 237th codon from AAG to AAA compared to the original *gfp-opt* sequence. 34 bp of the fructose-bisphosphate aldolase (FbaA) leader was included in the forward primer. The PCR fragment amplifying *gfp-opt*, was digested with PstI and HindIII and ligated into pMK1167/ $P_{rrnC}$ , resulting in the recombinant plasmid pAE28.

pAE28 was integrated in the chromosome of MG1363 by the use of the bacteriophage TP901-1 integration system (Brøndsted & Hammer, 1999). Utilizing the ability of the temperate bacteriophage TP901-1 a non-replicating integration vector can by site-specific chromosomal integration be integrated into the chromosome of *L. lactis* at the attachment site *attB*. The integrase enzyme (encoded by *orf1*) was donated in trans located on the plasmid pLB95. pLB95 was excreted after successful integration of the integration vector by means of its temperature sensitive origin.

## 2.5 Results

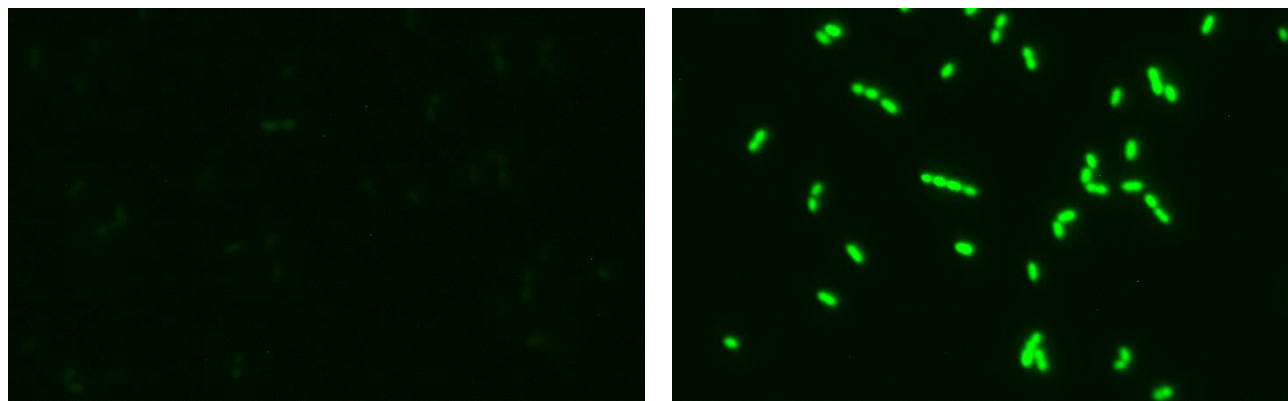
In *L. lactis* the green fluorescent protein has been expressed as a reporter for many different purposes; expression of an eukaryotic elongase, optimization of protein expression in fed-batch fermentations and test of novel promoter probe vectors (García-Cayuela *et al.*, 2012; Niu *et al.*, 2008; Oddone *et al.*, 2007). Common for most studies concerning GFP expression in *L. lactis* are that *gfp* is present on a multi-copy plasmid. To our knowledge only few studies have used chromosomal integration of *gfp* in *L. lactis*. Rawsthorne *et al.* used the group II intron, LI.ltrB, to achieve multicopy delivery of *gfp*, Scott *et al.* used transposon Tn916 to obtain a single chromosomal copy of *gfp* and Pinto *et al.* constructed a GFP containing integration plasmid (Pinto *et al.*, 2011; Rawsthorne *et al.*, 2006; Scott *et al.*, 2000).

In the study by Rawsthorne, Turner, and Mills 2006, as in the majority of other studies where GFP and *L. lactis* are combined, the *nisA* promoter is regulating expression of GFP. The strong inducible promoter

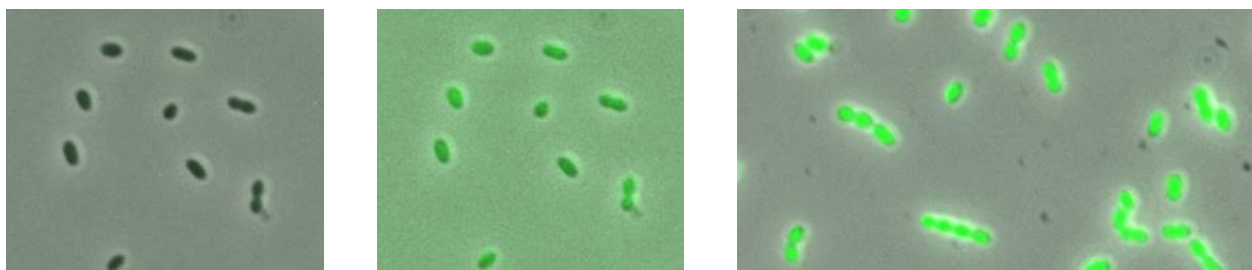
nisA is a part of the NICE expression system which is widely used (Oddone *et al.*, 2009). In the article from 2000 the nisA promoter turns out to be the only promoter, out of the three tested, which gives green fluorescent *L. lactis* cells (Geoffroy *et al.*, 2000).

We chose to integrate *gfp* into the chromosome, instead of expressing the protein from a multi-copy plasmid, to decrease the metabolic load on the cell and to avoid fluctuation in gene copy number. Fluctuation in gene copy number would potentially increase cell-to-cell variability and affect the conclusion drawn from the results. Regardless whether *gfp* is chromosomally integrated or plasmid borne a correlation between the copy number of *gfp* and the level of fluorescent GFP can be determined (Million-Weaver *et al.*, 2012; Rawsthorne *et al.*, 2006).

Creating a *L. lactis* reporter strains which was fluorescent enough for single cell analysis by the use of flow cytometry and which expressed GFP from a single chromosomally growth regulated copy of the gene required some preliminary attempts. The results of one of the preliminary strain constructions can be viewed in Figure 5(left). The same strain is shown in Figure 6(left) and Figure 6(middle). Both strains were growing balanced in chemically defined medium when the sample for microscopy was taken.



**Figure 5** Fluorescent microscopy images of *L. lactis* reporter strain AE064 (left) and *L. lactis* reporter strain AE072 (right). For both images was the exposure time 990 ms and the detector sensitivity (grain) was set to maximum level The filter set used to observe GFP expressing strain was a BrightLine® single-band filter set from Semrock with an excitation spectrum between 475 nm and 500 nm and an emission spectrum ranging from 510 nm to 540 nm.



**Figure 6** Fluorescent microscopy images of two *L. lactis* strains expressing GFP. (Left) Microscopy image of reporter strain AE064, the image is an overlay image of a fluorescent image and a phase contrast image. (Middle) same images as (left), but with modification in contrast and brightness, this image is included to show that the cells are in fact fluorescent. (Right) microscopy image of reporter strain AE072, this image is an overlay image of a fluorescent image and a phase contrast image, the settings used for this image are the same as for the image of AE064 show in (left).

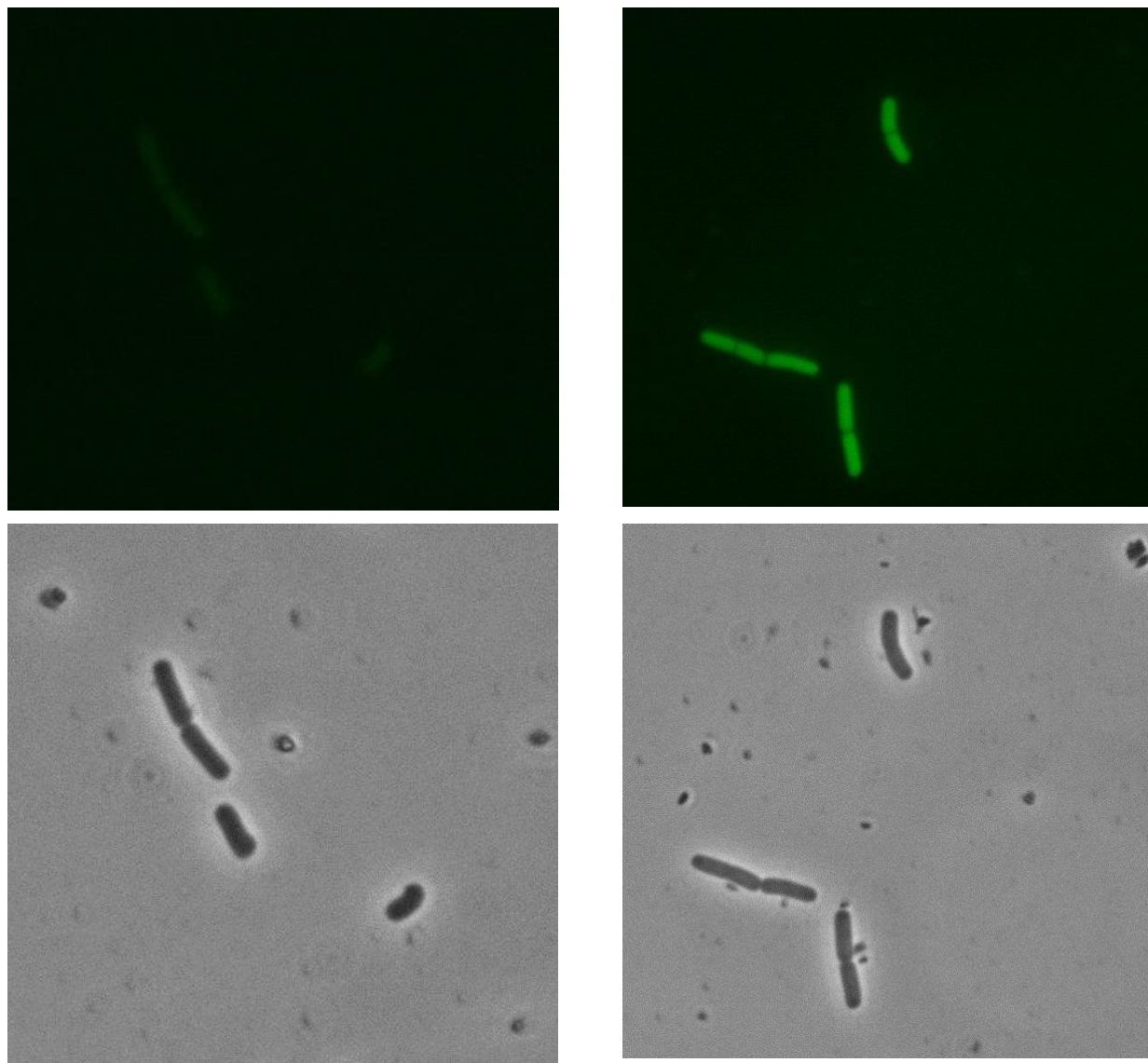
The ribosomal promoter sequence used in the two strains AE064 and AE072 was identical. The *L. lactis* fructose biphosphate aldolase (*fbaA*) leader was included in both reporter strains to increase the translation initiation. The only difference between the reporter strains AE064 and AE072 was the *gfp*. The choice of GFP variant and the *gfp* codon usage clearly had an effect on the level of fluorescence; AE064 was expressing a synthetically constructed *L. lactis* codon optimized genes (laGFP) which has the following mutation according to wtGFP: S65A/V68L/T203I. The strain AE072 was expressing *gfp-opt*, a codon optimized (for *Bacillus anthracis*) GFPmut1 (F64L/S65T) variant. It should be noted that *E. coli* cells, used for construction of the integration vectors, were bright fluorescent regardless of which of the vectors they was transformed with (data not shown).

Also in *B. subtilis* did the *gfp* codon usage, the variant of GFP and the sequence of the ribosomal binding site have a huge impact on the fluorescence intensity of the GFP reporter strain. To illustrate the difference in fluorescence intensity between two *B. subtilis* GFP reporter strains AE099 (which was used to analyze population heterogeneity by flow cytometry, see section 3 and section 6) and AE003, fluorescence microscopy images were taken. The settings (the detector sensitivity (gain), and the camera exposure time) used for the images of AE003 (Figure 7, top left) and AE099 (Figure 7, top right) were the same (gain 255 and 770 ms exposure time).

In both AE003 and AE099 *gfp* were present in one chromosomal copy and the same sequence of the 16S rRNA promoter  $P_{rrnA_1}$  was used for transcriptional regulation of the GFP expression in both strains. The recombinant plasmid pDG268neo was used for construction in both cases. In both strains were a strong ribosomal binding site placed upstream *gfp* to increase the translation initiation rate. In the reporter strain AE099 the ribosomal binding site was identical to the 40 bp upstream the start codon of *B. subtilis* strain 168 fructose biphosphate aldolase gene (*fbaA*) – the Shine Dalgarno sequence (3'-AAGGAGG-5') was here placed 9 bp upstream the start codon. In AE003 the Shine-Dalgarno sequence 3'-AAAGGAG-5' was placed

7 bp upstream the start codon. The GFP variant in AE099 was GFP-opt, a GFPmut1 variant codon optimized for expression in *Bacillus anthracis* (Sastalla *et al.*, 2009). The GFP variant expressed in AE003 was yGFP, a GFPmut2 variant codon optimized for *E. coli* (Hansen & Atlung, 2011).

Both in *L. lactis* and *B. subtilis* did differences in the design of the reporter strain, had a large effect on the fluorescence intensity of the cells.



**Figure 7** Fluorescence microscopy images of *B. subtilis* reporter strains AE003 (top left) and AE099 (top right). The settings (detector sensitivity (gain) and camera exposure time) were identical for the two fluorescent images (gain 255 and 770 ms exposure time). The phase contrast microscopy images (AE003, bottom left and AE099, bottom right) were included to show the location of the cells. The samples for the microcopy were taken from balanced growing cultures propagated in chemically defined medium supplemented with glucose.

### 3 Heterogeneity at a single cell level in homogenous cultures.

Anne Egholm Pedersen<sup>1</sup>, Camilla Thyregod<sup>2</sup>, Mogens Kilstrup<sup>1</sup>, and Jan Martinussen<sup>1\*</sup>

<sup>1</sup>: Metabolic Signaling and Regulation Group, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>2</sup>: Institut for Matematik og Computer Science, DTU compute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Running title: Heterogeneity in homogenous cultures

\*Corresponding author. Mailing address:

Metabolic Signaling and Regulation Group, Department of Systems Biology, Building 301,  
Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Phone: +45 45252498, E-mail: jma@bio.dtu.dk

Keywords: Population heterogeneity, GFP, flow cytometry, balanced growing cultures.

#### 3.1 Abstract

Traditionally, an isogenic culture in the exponential growth phase is considered to be homogenous. This means that the growth rates of all individual cells are assumed to be similar. We know today, based on published research, that an isogenic population of cells is heterogeneous. So to measure the actual distribution of growth rates at the single cell level, a *gfp* gene under the control of a growth rate dependent ribosomal RNA promoter, was integrated on the chromosome of the industrial relevant microorganism *Bacillus subtilis*. As expected, the expression of *gfp* at the population level was shown to be closely related to growth. In order to determine the distribution of growth rates at the single cell level, flow cytometry data were collected from growing cultures. To define the degree of heterogeneity statistical analysis of large data sets collected by flow cytometry, was central in characterizing the actual distribution of growth rates at a single cell level. A set of experiments addressed the importance of the inoculum culture history on the population heterogeneity in apparent homogenous cultures. Stationary and exponentially growing pre-cultures in both complex and chemically defined medium were used as inoculations in growth experiments. Growth rate heterogeneity was found to be increased when stationary phase pre-cultures were used as inoculum compared to the use of exponentially growing cultures in chemically defined medium. The study offers quantitative validating of the balanced growth dogma from the Copenhagen School of Bacterial Growth Physiology headed by Ole Maaløe in the fifties and sixties, which stated that a bacterial culture must have been exponentially growing, with the same growth rate, for at least eight generations, before valid and

reproducible physiological data can be obtained. Consequently the conclusions underline the importance of a proper experimental setup for studying cell physiology.

## 3.2 Introduction

Physiological diversity is an important issue when microorganisms are used as production organisms and uncontrolled increase in population heterogeneity may reduce the overall performance during a fermentation process and is thus undesirable in large scale fermentations. It has previously been shown that the use of a viable and optimized inoculum can improve fermentation productivity in large scale fermentations (Martínková *et al.*, 1991; Neves *et al.*, 2001). But even in isogenic optimized cultures phenotypic heterogeneity is a well-known fact (Elowitz *et al.*, 2002; McAdams & Arkin, 1997). In this study of population heterogeneity we chose to focus on the Gram positive organism; *B. subtilis*, which is an important production organism of industrial enzymes (van Dijk & Hecker, 2013; Westers *et al.*, 2004).

Outside the optimized environment of a fermenter or shake flask, feast and famine are inevitable conditions. Most natural niches are unable to sustain balanced growth, which is obtained when all cell components in a culture of isogenic cells increase exponentially with the same rate (Ingraham *et al.*, 1983; Schaechter, 2006). In nature, subpopulations with increased heterogeneity emerge in stressing environments. In a process called bet-hedging some of these apparently non-beneficial epigenetic hibernation states may increase the overall fitness of the species in case the environmental conditions suddenly should change (Slatkin, 1974). For use in controlled physiological experiments on the other hand heterogeneity within bacteria populations should be minimized to increase reproducibility of the experiments. Growth of bacterial cultures is performed as a standard technique on a daily basis in laboratories worldwide, and is the foundation on which the reproducibility of many types of experiments rely, including studies of physiology, expression of heterologous protein, transcriptomic and proteomic analyses, just to mention a few. All these experiments are performed on presumably homogenous cultures based on measurements of population averages.

At the population level Ole Maaløe and his colleges elegantly elucidated how the bacterial growth rate in balanced cultures correlate with the cell composition, and how the molecular cellular composition changes in up and down-shift experiments (Kjeldgaard *et al.*, 1958). Balanced growth for a batch culture is defined by a culture in steady state growing in a medium where all the relevant nutrients are present in excess. Because the growth rate is only limited by the absence of nutrients and not by their concentration the situation can be defined as 'unrestricted growth' (Schaechter *et al.*, 1958).

Based upon measurements of the cellular composition of balanced cultures, Schaechter *et al.* concluded that no matter how the medium was altered, the average composition of the cell correlated with the growth rate and they stated that: "Failure to maintain balanced growth throughout an experiment makes it impossible to relate any measured quantity to the growth rate in a direct way" (Schaechter *et al.*, 1958). Kjeldgaard *et al.* clearly illustrated how a balanced bacteria culture undergoes differential changes in the growth rate, the RNA content, the DNA content, the rate of protein synthesis and the cell division after a shift in nutrient status (Kjeldgaard *et al.*, 1958), leading to cells with vastly changing composition. In an ideal state of balanced

growth all cell components must, by definition increase at same exponentially rate (Ingraham *et al.*, 1983). The specific growth rate  $\mu$  is usually calculated from the increase in optical density (OD) of the medium from the equation  $dN/dT = \mu N$ , which is equal to  $\mu = \ln 2 / (\text{doubling time})$ . During balanced growth the increase in any cellular component (X) follows the equation  $dX/dT = kX$ , where k is equal to  $\mu$ . The growth rate is linked to the activity of ribosomal promoters, observable by the linear correlation between ribosome content, cell size, and growth rate (Ecker & Schaechter, 1963; Neidhardt & Magasanik, 1960; Schaechter *et al.*, 1958). Ribosomal promoters expressing a reporter gene can therefore be used to measure growth rate, as previously shown (Miura *et al.*, 1981; Sternberg *et al.*, 1999).

Green fluorescent protein (GFP) is a widely used reporter which emits green light when excited with blue light in the presence of  $O_2$ . It can be monitored non-invasively in living cells by fluorescence microscopy and flow cytometry (Cody *et al.*, 1993; March *et al.*, 2003). GFP has been validated against other reporters (LacZ and CAT) and has been demonstrated to be equally reliable for quantification of promoter activity (Albano *et al.*, 1998; Scholz *et al.*, 2000; Silva-Rocha & de Lorenzo, 2012). Distribution of GFP expression among individual cells can be determined with the powerful single cell analysis method of flow cytometry.

In an adjacent paper, we have considered different statistical methods to describe heterogeneity at the single cell level (Thyregod *et al.*, 2015). Here we suggested guidelines for a non-statistician working with flow cytometry on how to analyze data. The statistical analyses are made using the statistical point-and-click software JMP® version 11.

In the current study, we utilize this statistical tool to address how the history of the inoculum affects the population heterogeneity during batch cultivations in shake flasks. Especially the use of stationary cultures as inoculum was important as the entire metabolism is rewired when cultures enter the stationary phase. In our experiments the growth rate was determined by the composition of the medium and not by limitation of components, as being the case in chemostat cultivations. We show that the heterogeneity in an exponentially growing culture increases profoundly if the pre-culture used as inoculum had entered stationary phase.



### 3.3 Materials and Methods

**Table 4 Strains.**

Strain	Relevant characteristic	Source of reference
<i>E. coli</i> DH5α	F80/ <i>lacZ</i> DM15 D( <i>lacZ</i> YA- <i>argF</i> )U169 <i>recA1 endA1 hsdR17 supE44 thi-1 gyrA96 relA1</i>	Laboratory strain
<i>B. subtilis</i> strain 168	<i>trpC2</i>	(Burkholder & Giles, 1947; Spizizen, 1958)
AE044	<i>B. subtilis</i> strain 168 <i>amyE</i> ::pDG268neo	This study
AE099	<i>B. subtilis</i> strain 168 <i>amyE</i> ::pAE30	This study

**Table 5 Plasmid used for strain constructions.**

Plasmid	Relevant characteristic	Source of reference
pDG268neo	Ap <sup>r</sup> Nm <sup>r</sup> ; a vector for constructing transcriptional <i>lacZ</i> fusions designed to integrate in the chromosomal <i>amyE</i> gene which contains a promoterless <i>lacZ</i> gene and neo resistance cassette in opposite directions flanked by the truncated <i>amyE</i> gene	C. W. Price, Department of Food Science and Technology, University of California, Davis, California, USA.
pAE30	<i>gfp-opt</i> from (Sastalla <i>et al.</i> , 2009) in <i>Bam</i> HI: <i>Pci</i> I. P <sub><i>rrnA</i></sub> in <i>Eco</i> RI: <i>Bam</i> HI	This study
pSW4-GFPopt	Vector containing <i>gfp-opt</i>	(Sastalla <i>et al.</i> , 2009)

**Table 6 Oligonucleotides used in this study**

Primer	Sequence (5'-3')	Description
<i>rrnA</i> _fwd	GCTAGAATTCGTTACAACAGCTATCAGCGG	<i>rrnA</i> for pAE30
<i>rrnA</i> _rev	GCTAGGATCCCCAGTCTTACAGGCAGGTT	<i>rrnA</i> for pAE30
Fba( <i>Bam</i> HI)2_fwd	GCTACTGGATCCCGCTAGAAGACAATCAGGCTACATAAGGAGGACATTCGTCATGTCAAAAGGTGAAGAATT	<i>gfp-opt</i> for pAE30
IaGFP( <i>Pci</i> I)_rev	CGAAGTACATGTTATTTATATAATTCATCC	<i>gfp-opt</i> for pAE30

**Strains and DNA manipulations.** In this study the *E. coli* strain DH5α was used for genetic manipulations (Table 1). Details of all primers and plasmids are found in tables 2 and 3.

The *E. coli* cells were transformed by electroporation as previously described (Sambrook and Russell, 2001). *E. coli* were propagated in lysogeny broth (LB) (Bertani, 1951), and selected on LB plates supplied with

ampicillin to 100 µg/ml. Transformation of *B. subtilis* was performed as described by (Boylan *et al.*, 1972), transformants were selected on LB plates supplemented with 5 µg/mL neomycin.

**Construction of pAE30 and chromosomally integration in *B. subtilis*.** *gfp-opt* was amplified from pSW4-GFPopt) using the primers Fba(BamHI)2\_fwd and laGFP(PciI)\_rev. A 40 bp fragment of the FbpA leader, including the strong ribosomal binding site present in the leader, was placed upstream *gfp-opt*. The resulting PCR fragment was digested with BamHI and PciI and ligated into the *B. subtilis* integrative vector pDG268neo, digested with the same restriction enzymes. The recombinant plasmid was digested with EcoRI and BamHI and ligated together with a PCR fragment made by amplifying the P<sub>rrnA</sub> promoter (16S rRNA promoter, including both P1 and P2 and 257 bp downstream +1) using the primers rrnA\_fwd and rrnA\_rev. The recombinant plasmid (pAE30) was amplified in *E. coli* DH5α. Plasmid pAE30 was integrated into the *B. subtilis* 168 *amyE* gene by a double cross-over recombination through the *amyE* N-terminal and *amyE* C-terminal parts located on the pDG268neo vector. The plasmid was linearized by digestion with XhoI prior to transformation in order to ensure chromosomal integration by a double cross-over event. The neomycin-resistant strain AE099 was obtained. The control strain AE044 was obtained by transforming *B. subtilis* strain 168 with XhoI-linearized pGD268neo.

**Media.** *B. subtilis* was grown in Spizizen salt medium supplemented with 1 mg of thiamine and 40 mg tryptophan per liter, as previously described (Saxild & Nygaard, 1987). This medium will be referred to as chemically defined medium. Carbon sources were added to a final concentration of 0.4 %. Glutamate, if added, was used in a concentration of 0.2 %. The strains were cultivated in Erlenmeyer flasks at 37° C with vigorous shaking.

**Fluorescence measurements.** Fluorescence measurements at culture levels were carried out with a Shimadzu RF-5301PC fluorometer (Columbia, MD). GFP excitation was performed at 488 nm, and emission was monitored at 510 nm with excitation and emission slit widths of 5:5. The fluorescence of *B. subtilis* reporter strain and control strain was determined directly in the chemically defined medium.

**Fluorescence microscopy.** Microscopy was done using a Zeiss Axioplan fluorescence microscope, GFP fluorescence was analyzed with the use of Semrock GFP-3035C BrightLine single-band filter set. Fluorescence microscopy was used to isolate transformants.

**Flow cytometry.** For flow cytometric analysis the cells were diluted in PBS solution: NaCl, 0.137 M; KCl, 3 mM; NaHPO<sub>4</sub>, 12 mM; and KH<sub>2</sub>PO<sub>4</sub>, 2 mM; pH 7.4, and measured directly on a BD FACSCalibur (Becton-Dickinson, NJ, USA) flow cytometer. Excitation wavelength for the laser used was 488 nm. Fluorescence emission levels were measured using a band pass filter at 530/30 nm (FITC). For each sample approximately 100,000 cells were analyzed. Contour plots for visualization of flow cytometry data were made in FlowJo. The contour plots used have 5 % probability contours, meaning that 5 % of the cells lie between each contour.

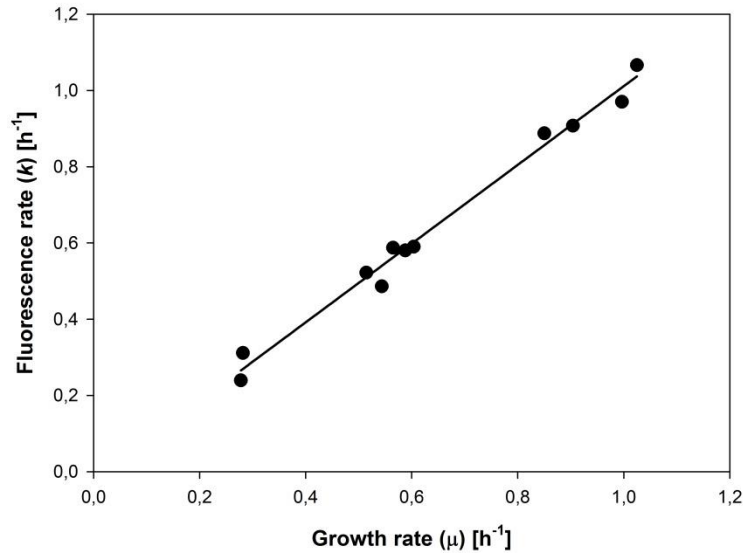
**Gating and setting forward scatter threshold before statistical analysis of flow cytometry data.** Data obtained from flow cytometry analysis was initially gated based on fluorescence; Flowing Software (v2.5.1, Turku Centre for Biotechnology, Turku, Finland) was used for histogram gating, removing background fluorescence defined by the control strain. The fluorescence per forward scatter (specific cellular fluorescence) was calculated for each event and the data was further analyzed in JMP®.

### 3.4 Results

#### 3.4.1 Expression of GFP from $P_{rrn}$ -gfp fusions correlate with the growth rate at the population level

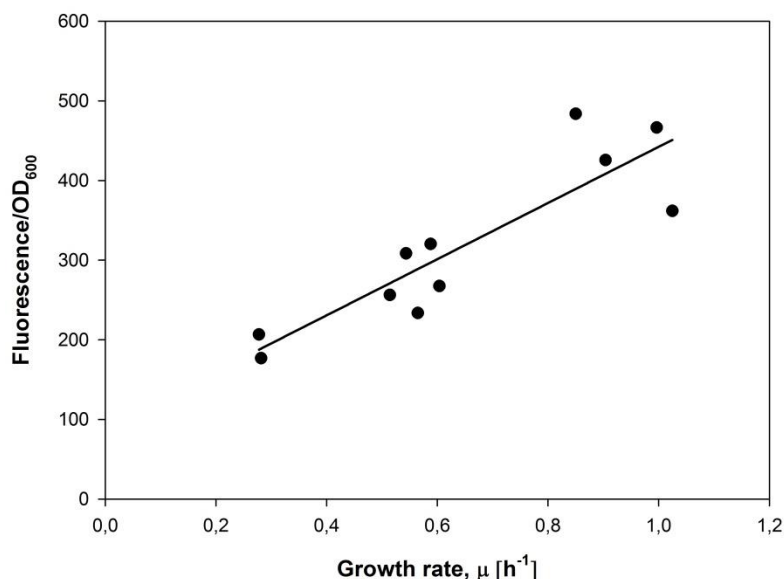
The most profound physiological parameter of a microbial culture is its growth rate, and it is the single parameter that best describes the average physiological state of the cells. In this study we wanted to address the very fundamental question, in which way the (macroscopic) growth rate of an exponential growing culture is determined by the (microscopic) growth rates of the individual bacteria. To assess the growth rate at both single cell and culture levels, a *B. subtilis* reporter strain expressing growth-rate regulated  $P_{rrn}$ -gfp fusion was constructed. The ribosomal content of a bacterial cell is correlated to its growth rate, by a fine-tuned regulation of ribosomal RNA transcription (Condon *et al.*, 1995; Gourse *et al.*, 1996; Samarrai *et al.*, 2011). Therefore, we chose to fuse a codon-optimized *gfp-mut1* gene (Cormack *et al.*, 1996; Sastalla *et al.*, 2009) to the *B. subtilis* ribosomal promoter *rrnA*. To reduce the metabolic load and to avoid copy number effects the *gfp* construct was integrated in a single copy on the chromosome of the host strain.

In an initial validation of the  $P_{rrn}$ -gfp expression as a function of the culture growth rate, different growth rates were obtained by changing the composition of chemically defined medium. In Fig. 1A the specific fluorescence rate  $k$  (obtained from the equation  $dFI/dT = k FI$ ) is plotted against the growth rate  $\mu$  to show how well balanced the cultures performed with respect to increase in OD and fluorescence. The strong linear relation shows that all cultures were in the balanced state.



**Fig. 1A Balanced growth and fluorescence. Relation between population growth rates ( $\mu$ ) and fluorescence rate ( $k$ ) for *B. subtilis* AE099 ( $P_{rm}\text{-gfp}$ ). The different growth rates and fluorescence accumulation rates were obtained in chemically defined medium with different carbon sources and with or without addition of glutamate.**

The level of fluorescence was correlated to the growth rate for each of the balanced cultures (Figure 1B). It shows a plot of the specific fluorescence (fluorescence/OD) against the culture growth rate. It is obvious that there is a linear correlation between the specific fluorescence and the growth rate, but a high fluctuation is observed despite the balanced state of all cultures. The fluctuation appears to increase with increasing growth rate and total fluorescence.

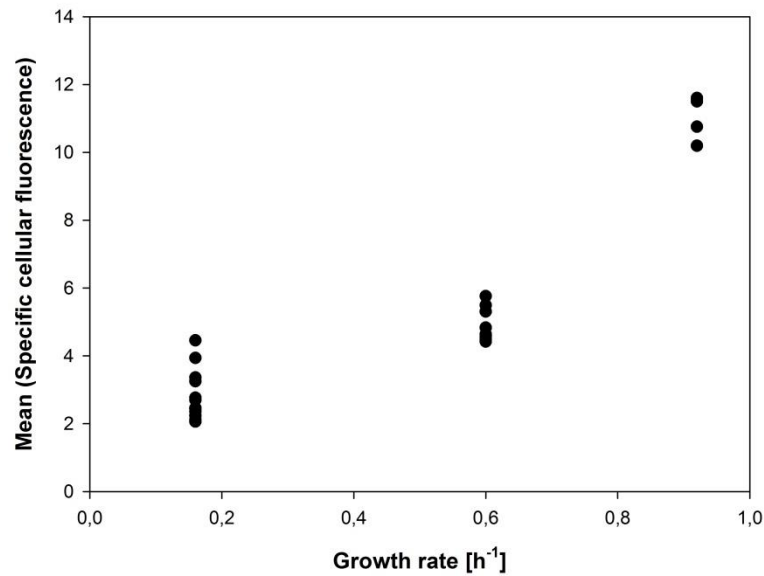


**Fig 1B. Correlation between growth rate and specific culture fluorescence.** Each data point originates from a plot of fluorescence versus OD<sub>600</sub> for one of the balanced cultures shown in panel A. The slope (fluorescence/OD) of the linear dependency is plotted against the specific growth rate of the culture

### 3.4.2 Distribution of individual growth rates within balanced cultures

Above we showed that the specific culture fluorescence from the *P<sub>rm</sub>-gfp* reporter strain, measured by a spectrofluorometer, correlated linearly with the growth rate. Using flow cytometry, fluorescence can also be determined at the single cell level. Just as the specific culture fluorescence can be defined as the total culture fluorescence divided by the biomass concentration (OD), a specific cellular fluorescence can be defined for a single bacterium as the value obtained by dividing the fluorescence signal by its forward scatter signal, as it represents cell size. Specific cellular fluorescence is constant when a cell divides, and represents the *P<sub>rm</sub>-gfp* expression per cell mass thus proportional to the growth rate of the individual cell.

To study the individual growth rates in media supporting different doubling times the *B. subtilis* reporter strain (AE099) was propagated at three different growth rates. The strain was balanced as before by growing exponentially in chemically defined medium for at least eight generations before flow cytometry analysis. Samples were withdrawn from the cultures during exponential growth, centrifuged and resuspended in ice cold PBS solution immediately before flow cytometry analysis. Figure 1C shows the correlation between the specific cellular fluorescence and the growth rate of a set of cultures. Due to stability issues of the flow cytometer the set of results had to be chosen from samples processed on the same day. We observed a strong correlation between growth rate and specific cellular fluorescence in accordance with the previous observed correlation between the growth rate and the specific culture fluorescence.



**Fig. 1C. Correlation between growth rate and specific cellular fluorescence. Each data point originates from a plot of fluorescence versus forward scatter for 100000 individual cells, analyzed by flow cytometry of samples from a balanced culture. The mean for each sample is plotted against the specific growth rate of the culture.**

The high fluctuation in specific cellular and specific culture fluorescence was not expected since all cultures grew balanced. To investigate the source of this variation we subjected the results to a descriptive statistical analysis. In Fig. 2 the specific cellular fluorescence of individual cells from three *B. subtilis* cultures are visualized as box-plots. The bottom and top of the boxes represent the first and third quartile and the line through the box represents the median value. The box plus whiskers include all data from the 2.5% percentile to the 97.5% percentile, i.e., 95% of all data points.

Box-plots provided us with an intuitive overview of the heterogeneity within a sample and between samples (Fig. 2). As concluded in our adjacent study (Thyregod et al., 2015) descriptive statistics offer a more detailed understanding of the underlying population heterogeneity. Figure 3 shows for each culture the sample mean, the standard deviation, the coefficient of variation (CV) and the interquartile range for the specific cellular fluorescence.

The visual overview provided in the box-plots (Fig. 2) together with the descriptive statistics (Fig. 3), shows a high degree of heterogeneity in specific cellular fluorescence, in a balanced culture that previously has been considered homogenous. Results from the slow growing culture are shown in Fig. 2A. It is evident from the growth curve (insert) that the culture grew exponentially, but the change in the level of specific cellular fluorescence suggests that the culture experienced a gradual lowering of overall growth rate until Time = 300 min. The variance of the signals appear to be highest in the start, but the coefficient of variation (relative standard deviation) is actually constant (Fig. 3A). The slow culture did not reach stationary phase during the experiment, but this was the case for the faster growing strain shown in Fig2B. Here it is apparent that the

specific cellular fluorescence is constant throughout the exponential growth but rises when the cell enters the stationary phase. This result does not agree with the strong correlation between growth rate and specific cellular fluorescence, and it is actually an artefact relating to the use of reporter fusions. When a bacterium enters stationary phase, its size lowers and it continues production of GFP even though the cells have stopped dividing. This has the consequence that the specific cellular fluorescence rises and leads to fluorescence levels that are detached from the activity of the promoter in front of the *gfp* gene. As a consequence of this artefact the correlation between growth rate and specific cellular fluorescence shown in Fig. 1C, was only found with data from balanced or nearly balanced cultures. At the highest growth rate (Fig. 2C) we observed a more constant fluorescence level, although the variance dropped considerable in the late exponential phase. We have no explanation for this behavior. Overall (Fig. 2A,B,C) the coefficient of variation was highest at the lowest growth rates showing that the growth rate distribution was much wider at lower rates. The median values of the specific cellular fluorescence during the exponential phases follow the expected correlation with the growth rates.

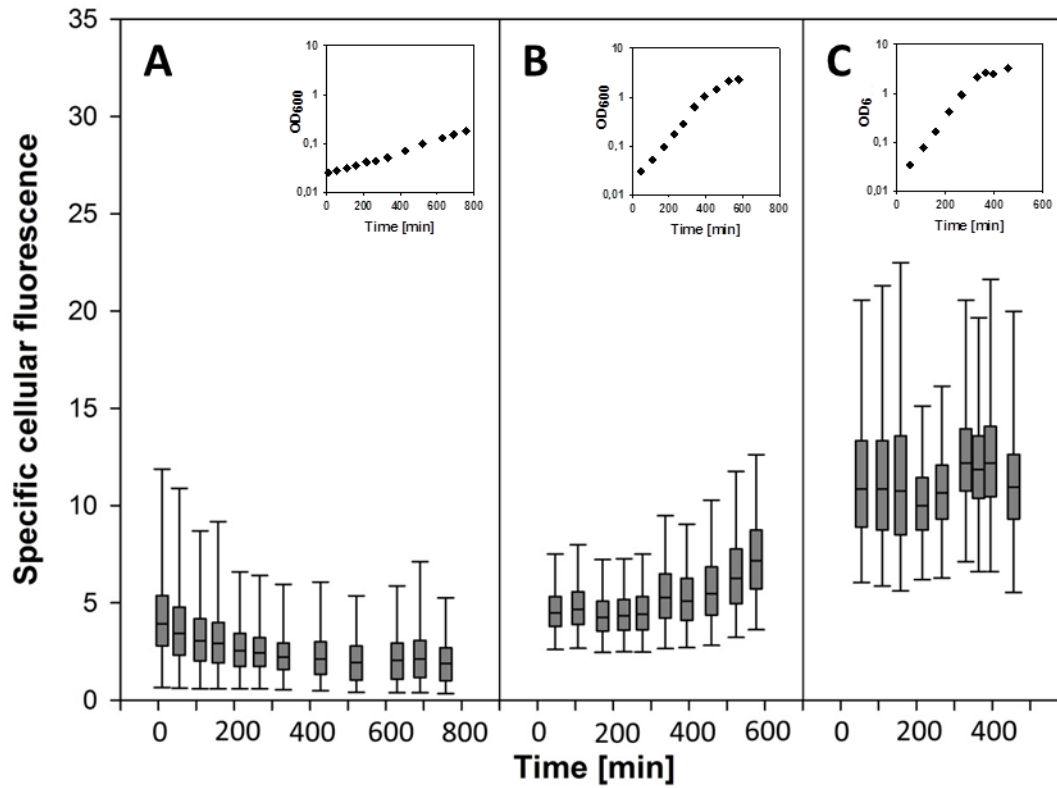
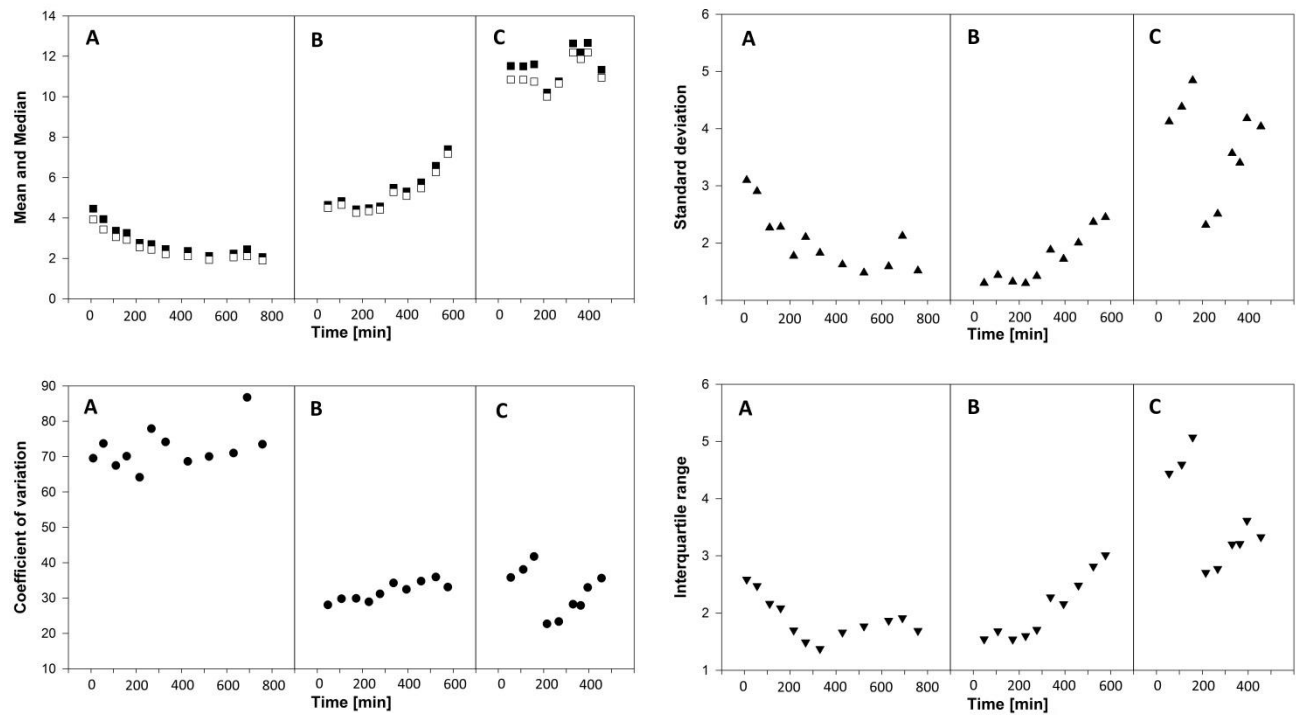
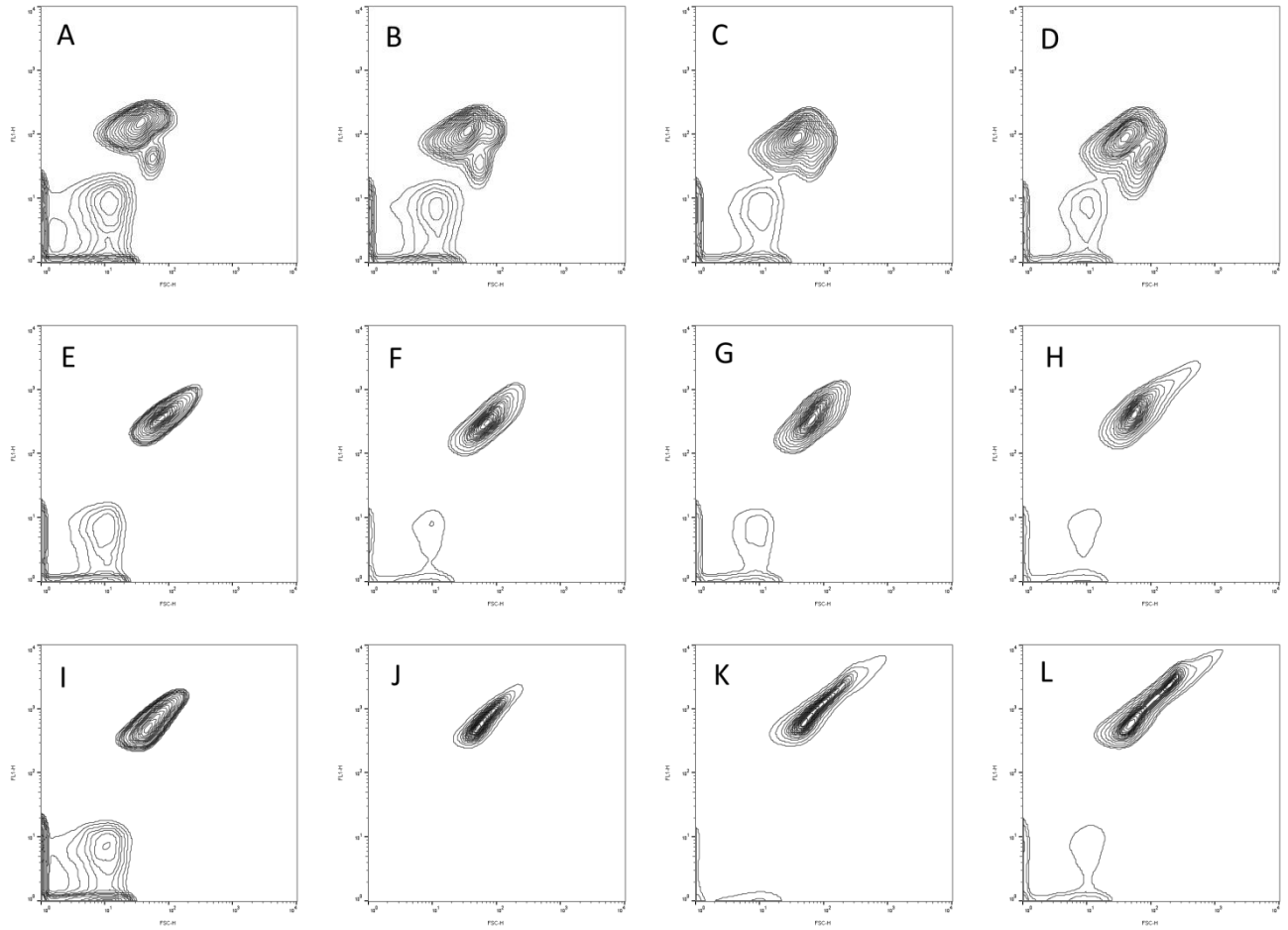


Fig. 2 Distribution of specific cellular fluorescence in cultures of *B. subtilis*, harboring a  $P_{rmaA}$ -gfp fusion, at three different growth rates presented in box-plots. The results are obtained by flow cytometry analysis of the GFP reporter strain. The points in the growth curve shown above each histogram indicate when the samples for flow cytometry were harvested. (A) *B. subtilis* GFP reporter strain cultivated on ribose (doubling time 256 min,  $\mu = 0.16 \text{ h}^{-1}$ ). (B) *B. subtilis* GFP reporter strain cultivated on glucose (doubling time 67 min,  $\mu = 0.62 \text{ h}^{-1}$ ). (C) *B. subtilis* GFP reporter strain cultivated on glucose and glutamate (doubling time 45 min,  $\mu = 0.92 \text{ h}^{-1}$ ).





**Fig. 3** Descriptive statistics describing specific cellular fluorescence computed from flow cytometry analysis of *B. subtilis* AE099 at three different growth rates. (A) *B. subtilis* GFP reporter strain cultivated on ribose (doubling time 256 min). (B) *B. subtilis* GFP reporter strain cultivated on glucose (doubling time 67 min). (C) *B. subtilis* GFP reporter strain cultivated on glucose and glutamate (doubling time 45 min). (Top left) Mean (black squares) and median (white squares). (Top right) standard deviation (triangles). (Bottom left) coefficient of variation (dots). (Bottom right) interquartile range (triangle, down). Each symbol can be linked to a point in the growth curves shown above the box-plots in Fig. 2.

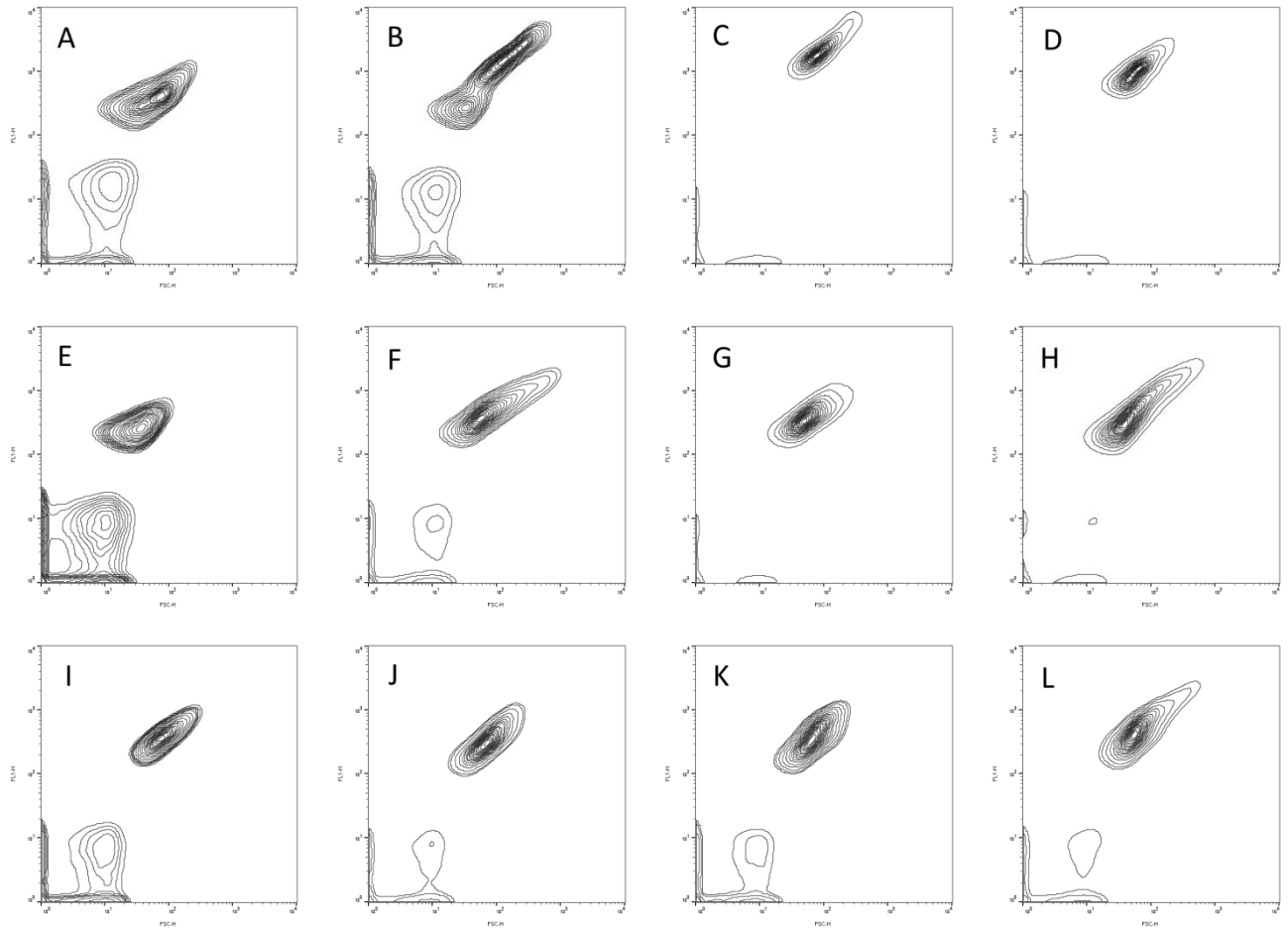


**Fig. 4** Contour plots for the *B. subtilis* reporter strain cultivated at three different growth rates. Contour plots with 5 % probability contours. A-D show distribution of forward scatter ([FSC-H], horizontal axis, log scale four decades) and fluorescence ([FL1-H], vertical axis, log scale four decades) for the *B. subtilis* culture propagated in chemically defined medium supplemented with ribose 10 min, 110 min, 428 min and 758 minutes after inoculation, see growth curve in Fig. 2 (a). Contour plots E-H shows forward scatter and fluorescence measured for a *B. subtilis* chemically defined glucose medium culture (doubling time 67 minutes) at 47 min, 228 min, 337 min and 577 min after inoculation (growth curve in Fig. 2 (b)). In contour plots I-L can fluorescence and size distribution for the *B. subtilis* GFP reporter strain cultivated as a balanced culture in chemically defined medium supplemented with glucose and glutamate 55 min, 158 min, 267 min and 395 min after inoculation be seen. The associated growth curve can be seen in Fig. 2 (c).

### 3.4.3 The level of heterogeneity is determined by the history of the inoculate

There are many ways for a researcher to prepare cell cultures for physiological investigations. In numerous studies, a stationary culture propagated in complex medium is used as inoculum, added at less than hundred-fold dilution. Based on the high variation in specific cellular fluorescence during balanced growth, we wanted to analyze to what extent the physiological condition of the pre-culture affects the heterogeneity of the culture. An increased cell-to-cell heterogeneity could have severe implications for the overall conclusions from an experiment. The experiments were designed as follows: *Stationary LB inoculum*: A pre-culture of *B. subtilis* was made by inoculation of a single colony from a LB agar plate into liquid LB. This pre-culture in broth was left to grow overnight into stationary phase and then diluted 150-fold into pre-warmed LB (Fig. 6A) and chemically defined glucose medium (Fig. 6B), respectively. *Balanced inoculum control*: A balanced *B. subtilis* culture was made by inoculating a single colony from an LB plate into chemically defined medium supplemented with glucose and propagated for approximately seven hours. This exponentially growing pre-culture was serially diluted in pre-heated chemically defined medium and grown overnight to serve as pre-cultures. Next morning a culture in the early exponential growth phase was used as inoculum and diluted into pre-heated chemically defined medium. The procedure ensured that the cells have been growing exponential for more than 20 generations. Samples from each of the three conditions were taken for biomass ( $OD_{600}$ ) determination and flow cytometry (Fig. 6C).

After inoculation of LB broth with a stationary phase LB culture, exponential growth was observed after a 90 minute lag phase. The lag phase before entering exponential growth was much longer (200 min.) when chemically defined medium was inoculated with the same stationary phase culture. As expected, inoculation with the exponentially growing pre-culture continued to grow exponentially. Contour plot of fluorescence signal per cell versus the forward scatter signal from samples of 100000 cells are shown in Fig. 5. Samples from four different time points are shown for each of the three experiments. It is clear from Fig. 5 that both the cell size distribution (x-axis showing forward scatter) and the fluorescence distribution changed continuously during the exponential phase following inoculation of the stationary LB culture into LB (Fig. 5A to D). The LB culture was clearly divided into subpopulations when it started to grow exponentially. Stationary cells are small and show low fluorescence (Fig. 5D). A homogenous population of such stationary phase cells (Fig. 5A) is seen gradually to develop into larger more fluorescent cells (Fig. 5B) which dominated during the exponential growth phase (Fig. 5C). This is in accord with previous knowledge that fast growing *B. subtilis* cells are larger than slow growing (Sargent, 1975; Sharpe *et al.*, 1998). This phenomenon is not observed in chemically defined medium after inoculation with a stationary LB pre-culture (Fig. 5E to H).



**Fig. 5. Contour plots, with 5 % probability contours, of *B. subtilis* cultures inoculated with pre-cultures at different physiological status. Horizontal axis: Forward scatter (FSC-H), vertical axis: GFP fluorescence (FL1-H), both log scale with four decades. A-D show *B. subtilis* reporter strain AE099 cultivated in LB, inoculated with a stressed LB overnight culture. The samples presented as contour plots were taken 55 min (lag phase), 108 min (exponential phase), 215 min (exponential phase) and 363 min (stationary phase) after inoculation, see growth curve in Fig. 6 (a). E-H show *B. subtilis* reporter strain AE099 cultivated in chemically defined medium supplemented with glucose, inoculated with a stressed LB overnight culture. The samples presented as contour plots were taken 57 min (lag phase), 267 min (exponential phase), 430 min (exponential phase) and 692 min (stationary phase) after inoculation, see growth curve in Fig. 6 (b). I-L show *B. subtilis* reporter strain AE099 cultivated in chemically defined medium supplemented with glucose, inoculated with an exponentially growing culture propagated in chemically defined medium supplemented with glucose. The samples presented in the contour plot were taken 47 min (exponential phase), 228 min (exponential phase), 337 min (exponential phase) and 577 min (stationary phase) after inoculation, see growth curve in Fig. 6 (c).**

In order to quantify the data from the flow cytometry experiments, they were analyzed and presented in box-plots (Fig. 6). Even though the LB culture (Fig. 6A) was exponentially growing and appeared to be in steady state from Time = 50 min, the sample mean of the specific cellular fluorescence changed dramatically from sample to sample (Figure 7A, top left).

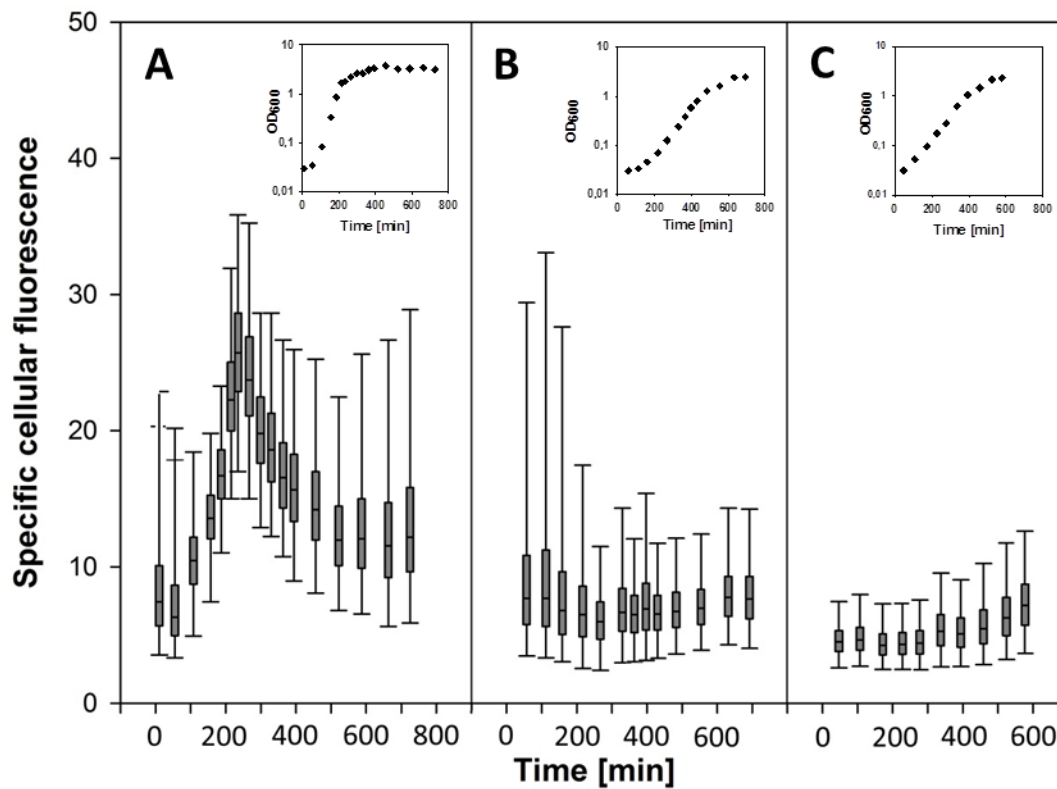
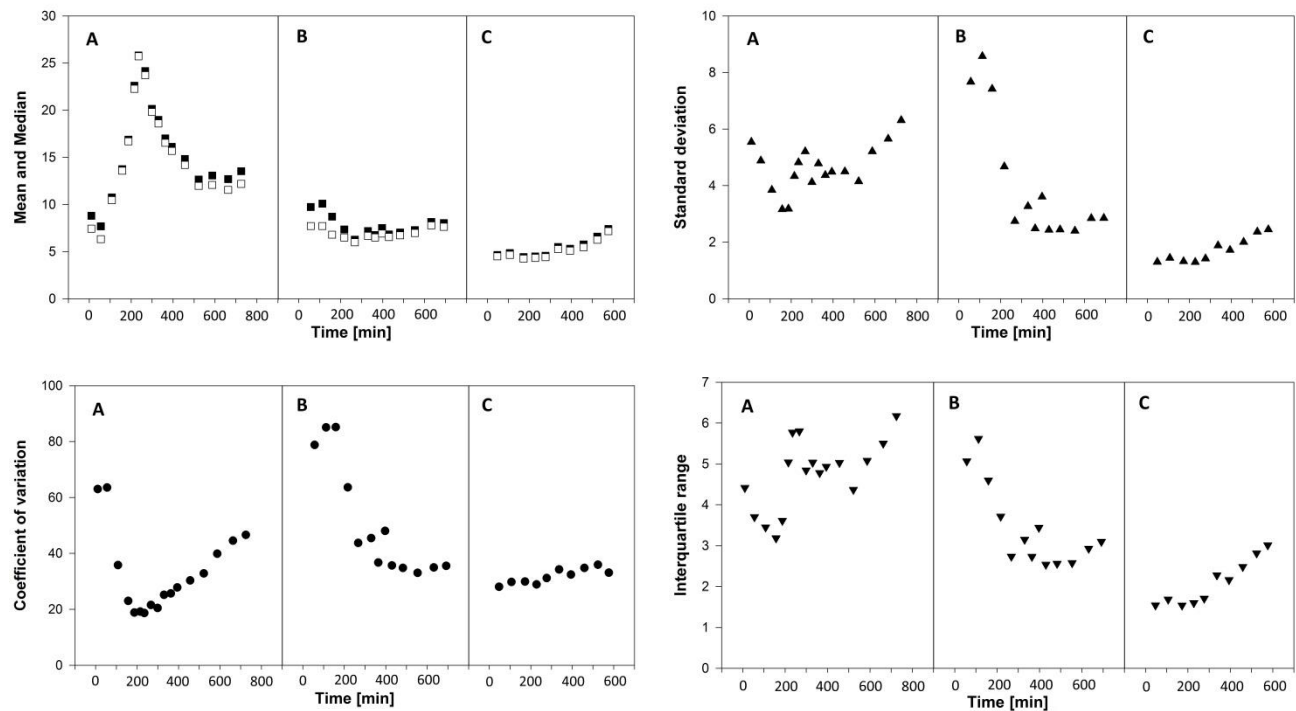


Fig. 6. Box-plot showing the distribution of specific cellular fluorescence for *B. subtilis* GFP reporter strain at three different growth conditions. The results were obtained by flow cytometry analysis of GFP reporter strain *B. subtilis* AE099. The points in the growth curve shown above each histogram indicate when the samples for flow cytometry were harvested. (A) Stationary phase pre-culture propagated in LB was used as inoculum in LB. (B) Stationary phase LB pre-culture was inoculated in chemically defined glucose medium. (C) An exponentially growing pre-culture propagated in chemically defined glucose medium was inoculated in chemically defined glucose medium (a balanced growing culture).



**Fig. 7. Descriptive statistics describing specific cellular fluorescence computed from flow cytometry analysis of *B. subtilis* AE099 in different growth conditions. (A) Stationary phase LB pre-culture of the *B. subtilis* GFP reporter strain was used as inoculum in LB. (B) Stationary phase LB pre-culture was inoculated in chemically defined glucose medium. (C) An exponentially growing pre-culture of the reporter strain was propagated in chemically defined glucose medium and inoculated in chemically defined glucose medium (a balanced growing culture). (Top left) Mean (black squares) and median (white squares). (Top right) standard deviation (triangles). (Bottom left) coefficient of variation (dots). (Bottom right) interquartile range (triangle, down). Each symbol can be linked to a point in the growth curves shown above the box-plots in Fig. 6**

As expected from the correlation between growth rate and specific cellular fluorescence, the cells were more fluorescent when propagated in LB compared to cultures in chemically defined glucose medium (Fig. 7A versus B and C, top left) due to the increased growth rate. The coefficient of variation for the LB culture continuously decreased during the exponential phase (Fig. 7A, bottom left panel) and appeared to be inversely correlated to the fluorescence level. The standard deviation decreased slightly during the exponential phase. The difference in mean value between the samples was very pronounced, leading to large sample to sample variation, the population profiles shown in contour plots, Fig. 5 (b-d), support this observation. It intuitively makes sense that the heterogeneity is reduced during the exponential phase when an un-balanced inoculum is used.

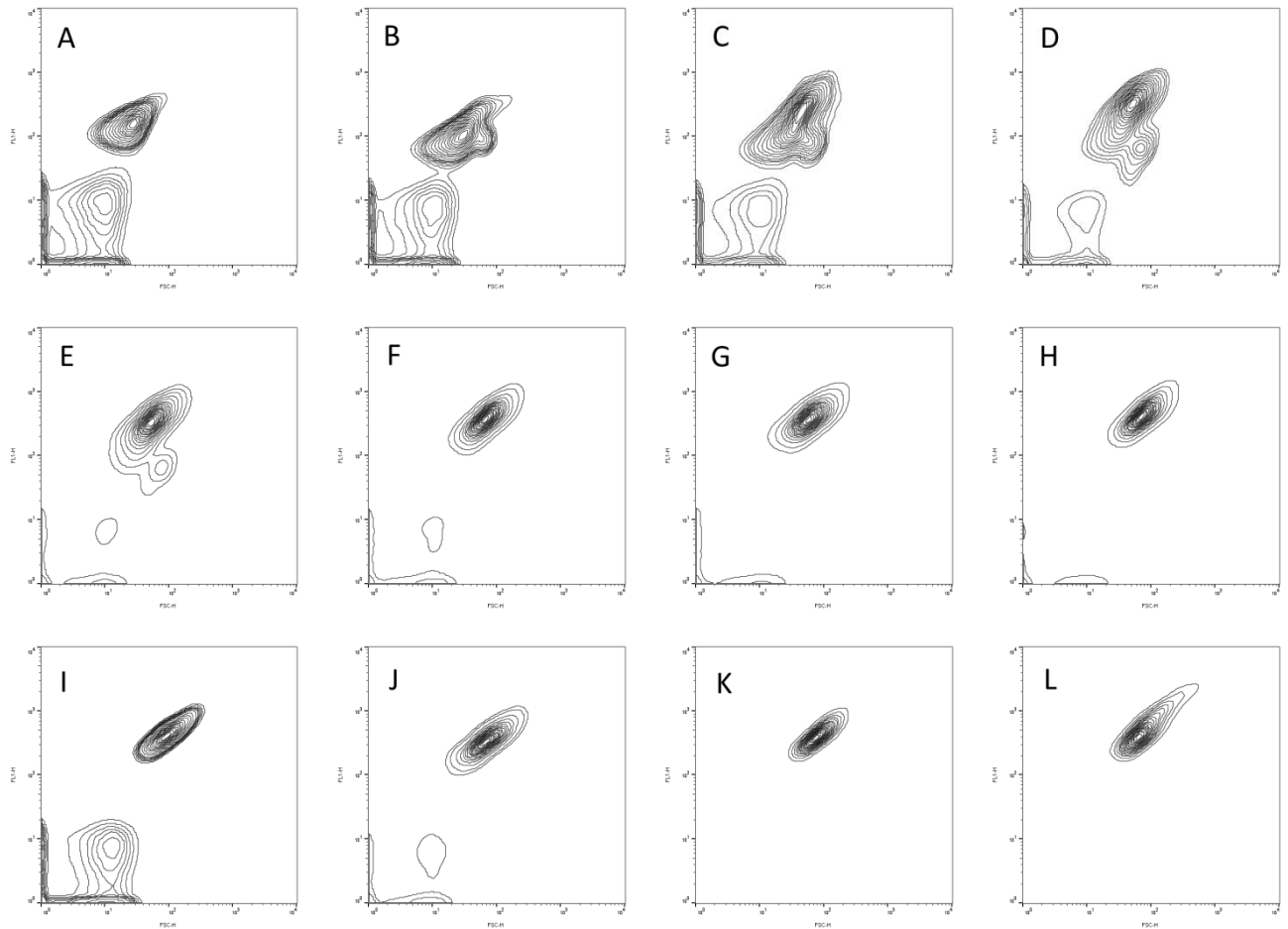
The mean and median fluorescence did not change much during the exponential phase for the cultures propagated in chemically defined medium (Fig. 7B and C, top left panel) whether the inoculum was from a stationary LB culture or a balanced culture in the same medium. This was somewhat unexpected, but a relative large heterogeneity, defined both by the interquartile range, the coefficient of variation and the length

of the box-plot whiskers, was observed during the lag phase and in the start of the exponential phase, when it was inoculated with the stationary LB pre-culture (Fig. 6B and Fig. 7B). In the late exponential phase the level of heterogeneity approached the low level observed for the balanced growing culture (Fig. 6C and Fig. 7C).

#### **3.4.4 Copenhagen School of Bacterial Growth Physiology revisited**

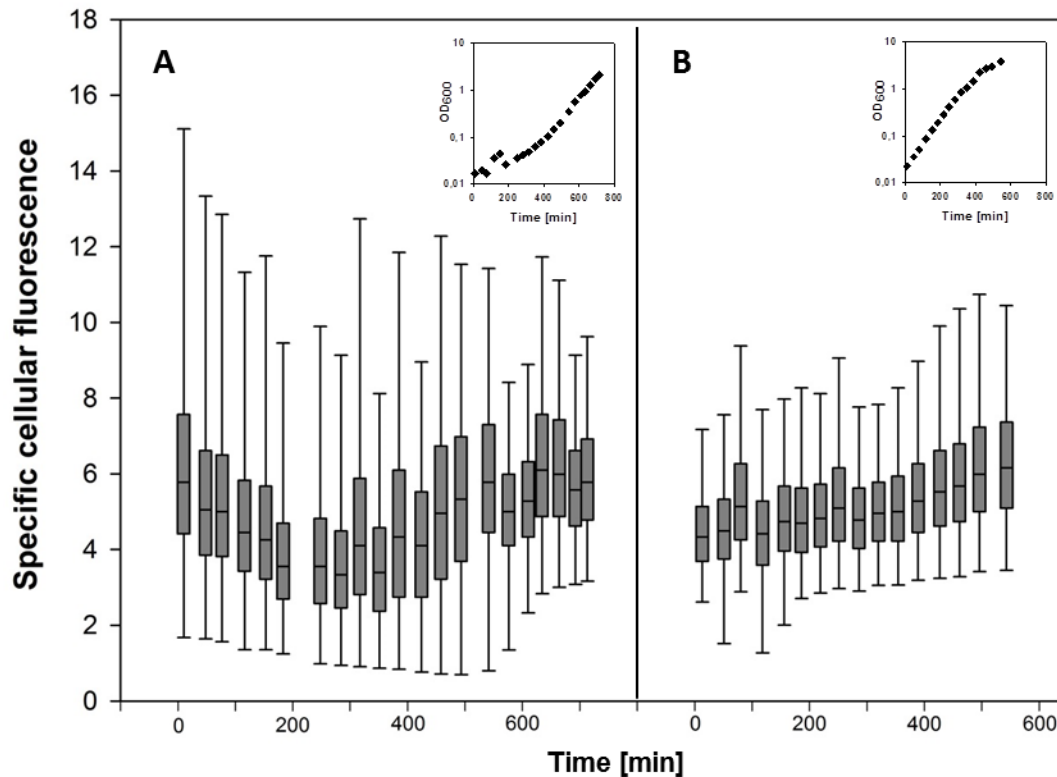
As emphasized in the introduction one of the dogmas of Copenhagen School of Bacterial Growth Physiology stated that a bacterial culture needs to grow exponentially for many generations to attain balanced growth (Schaechter, 2006), before reliable data can be extracted. Our results corroborate this statement and show that even though growth curves in terms of biomass production indicate a stable exponential growth condition, the growth rates of the individual cells are much more heterogenic in chemically defined medium when inoculated from stationary LB cultures than from exponentially growing culture propagated in chemically defined medium.

Inoculation of a culture in chemically defined medium with a stationary broth culture may not be common practice in most laboratories. However the use of outgrown cultures in chemically defined medium for inoculation of cultures in the same medium is common practice. To investigate i) whether stationary phase pre-cultures in chemically defined medium would influence population heterogeneity and ii) if propagations in several generations in fresh medium could minimize heterogeneity to the level of a balanced culture; a stationary phase pre-culture in chemically defined medium were used to inoculate chemically defined medium. Like previously, a balanced culture was included as control.



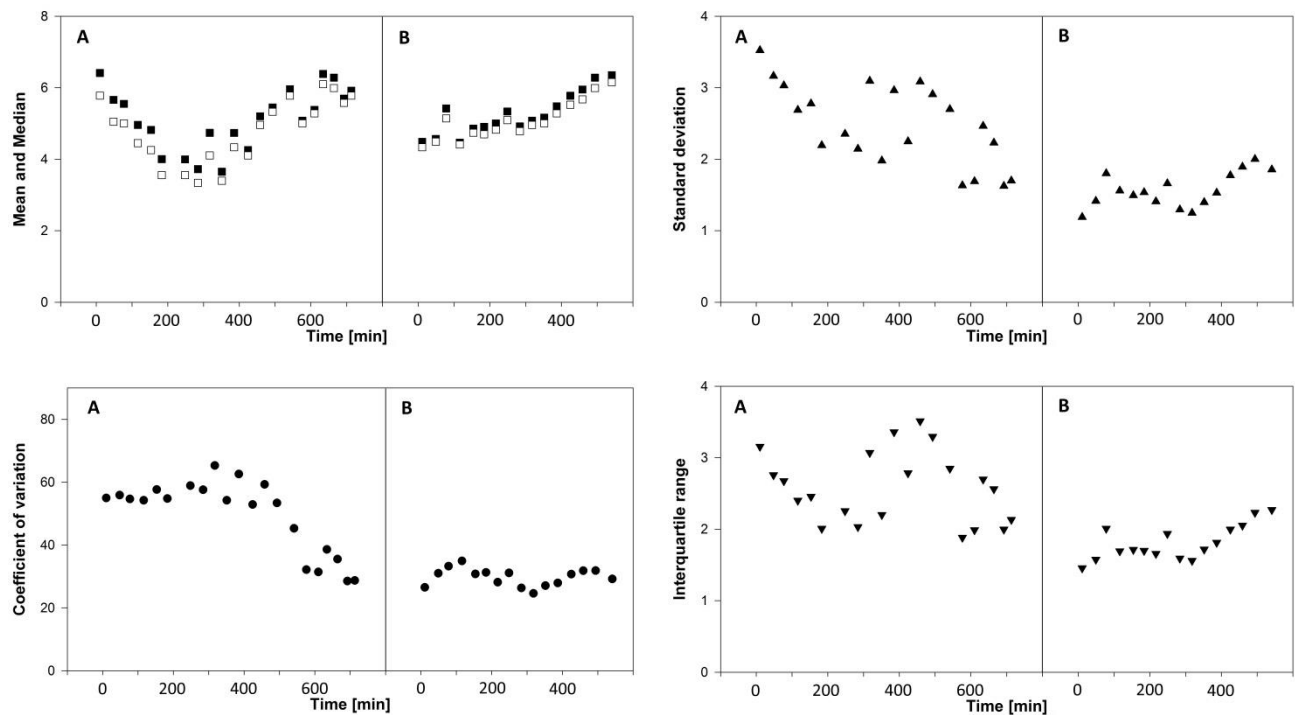
**Fig. 8.** Contour plots, with 5 % probability contours, displaying the fluorescence (horizontal axis) and forward scatter (vertical axis) for the *B. subtilis* GFP reporter strain inoculated with either a stressed stationary phase pre-culture (A-H) or a balanced exponentially growing pre-culture (I-L). The growth curves of the cultures can be seen in Fig. 9. Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose. Contour plots A-H display the development of size and fluorescent distribution for a culture inoculated with a stationary phase pre-culture A: 10 minutes after inoculation. B: 248 minutes after inoculation C: 385 minutes after inoculation. The culture had, according to optical density measurements, just started to grow exponentially at this time. D: 493 minutes after inoculation, 1.34 generations in exponential phase. E: 541 minutes after inoculation, 2.17 generations in exponential phase. F: 610 minutes after inoculation, 3.32 generations in exponentially phase. G: 634 minutes after inoculation, 3.58 generations in exponentially phase. H: 692 minutes after inoculation, 4.45 generations in exponentially phase. Contour plots I-L display the development of size and fluorescent distribution for a balanced growing culture I: Population profile just after inoculation, the culture started to grow exponentially immediately after inoculation. J: 184 min after inoculation (3.05 generations in exponentially growth). K: 318 minutes after inoculation (5.51 generations in exponential growth). L: According to optical density measurements the culture had, at the time this sample was taken, been in stationary phase for more than 100 minutes (sample taken 541 minutes after inoculation).





**Fig. 9. Distribution of specific cellular fluorescence for cultures of the *B. subtilis* reporter strain AE099 cultivated in chemically defined medium. The points in the growth curve shown above each histogram indicate when the samples for flow cytometry were harvested. The inoculum used was (A) a stationary phase chemically defined medium pre-culture and (B) an exponentially chemically defined medium pre-culture.**

The distributions of fluorescence and forward scatter for the culture inoculated with the stationary pre-culture are shown as contour plots in Fig. 8A to 8H. Accordingly, Fig. 8I to 8L show the distributions for the balanced culture. While the balanced culture show a constant distribution of specific cellular fluorescence versus size throughout the experiment (Fig. 8I to 8K), the inoculation with a stationary phase culture result in a change from small low-fluorescence cells to larger high-fluorescence cells (Fig. 8A to 8H). Even though both cultures according to cell density measurements appear to be growing exponentially, there is a large difference in their fluorescence and size distributions (compare Fig. 8C to 8H with Fig. 8I to 8K). After more than three generations (600 minutes after inoculation) the variation of the culture inoculated with the outgrown culture decreases to the level of heterogeneity of the balanced culture. This can be seen when comparing standard deviation, coefficient of variation of interquartile range between the two cultivations (Fig. 10).



**Fig. 10. Descriptive statistics describing specific cellular fluorescence computed from flow cytometry analysis of *B. subtilis* AE099 cultivated in chemically defined medium inoculated with (A) a stationary phase chemically defined medium pre-culture (B) an exponentially growing chemically defined medium pre-culture. (Top left) Mean (black squares) and median (white squares). (Top right) standard deviation (triangles). (Bottom left) coefficient of variation (dots). (Bottom right) interquartile range (triangle, down). Each symbol can be linked to a point in the growth curves shown above the box-plots in Fig. 9**

When visualized in a box-plot (Fig. 9) and the appertaining descriptive statistics (Fig. 10), the temporal succession of variation in specific cellular fluorescence supports the conclusions from the contour plots – the level of heterogeneity in a balanced growing culture was dramatically lower compared to level observed in the culture inoculated with a stationary phase pre-culture. Sample distributions were compared using qq-plots that were made by plotting the quantiles of two samples against each other and observing the fitted line, as extensively elaborated on in an adjacent study (Thyregod et al., 2015). A clear tendency was observed: The culture in steady state of balanced growth had, with only slight variation, a similar distribution of single cell growth rate throughout the exponential phase, this was not the case for the culture inoculated with a stationary phase pre-culture. Thus in order to decrease population heterogeneity and thus increase reproducibility, these results emphasize the use of balanced growing cultures.

### 3.5 Discussion

Flow cytometry offers great potential for the study of population heterogeneity, as it is possible to analyze more than 1000 events per second. Compared to fluorescence microscopy, flow cytometry thus makes it possible to collect large amounts of data, providing the statistical background for well-documented conclusions. By combining the powerful tool of flow cytometry with the reporter protein GFP, which have wide applications in visualizing bacterial gene expression and promoter activity, we were able to analyze heterogeneity of single cell growth rates. For evaluation of heterogeneity in cultures of isogenic bacteria reporter strains expressing growth-rate regulated GFP were made by integration of a  $P_{rm}$ -gfp reporter plasmid into the chromosomes of *B. subtilis*.

The fluorescence expressed by the individual cells is a combination of the rates of expression, GFP chromophore maturation, GFP degradation, and dilution of cellular GFP by cell growth and division. Due to the stability of GFP, a decrease in  $P_{rm}$ -gfp expression from a sudden lowering of the growth rate will not immediately be reflected in decreased fluorescence from that particular cell. Therefore, the single cell growth rates observed in this paper is not an instantaneous picture of the growth rate of the individual cell. Consequently, the heterogeneity may be underestimated during the shift situations. We are currently investigating whether an unstable GFP version would be better for assessing actual growth rates.

The reproducibility of physiological experiments, leading to firm conclusions, is highly dependent on a uniform physiological state of the culture. Stochastic variations appear as a result of heterogeneity at the cellular level, and an increase in heterogeneity within the population might result in reduced reproducibility. Variation between the samples, even though the level of heterogeneity in the population (within the sample) does not change, might also reduce the reproducibility of an experiment. In this work we have shown that a certain level of heterogeneity is inevitable, but that it can be reduced to a minimum by a proper experimental setup. Here we have shown that the physiological status of the pre-culture inoculum has a high influence on the level of heterogeneity. When stationary phase cells are used as an inoculum subpopulation of e.g. slower growing cells can be visualized using histogram plots of specific cellular fluorescence. It is obvious that the presence of subpopulation, during some of the exponential phase, would challenge the reproducibility of the experiments, as the results would vary dependent on sampling time.

When a culture of isogenic bacteria cells enters stationary phase, several morphological and physiological phenotypes change; cell volume, cell shape, compaction of the nucleoid, alteration in cell wall composition and accumulation of storage materials (Ishihama, 1997, 1999; Kolter *et al.*, 1993; Neidhardt *et al.*, 1990; Roszak & Colwell, 1987). Using a culture in stationary phase, where the cells have undergone these changes in phenotype, as inoculum for a culture in fresh medium, a reversal of the changes takes place during the lag phase as the cells prepare for exponential division. Every single regulatory event in the cells may potentially affect the phenotypic variation (Elowitz *et al.*, 2002), so the transition from stationary to exponential growth phase will increase the population heterogeneity as clearly demonstrated in this work.

The lowest level of heterogeneity in chemically defined medium cultures was observed in balanced growing cultures; a culture of isogenic cells, which has been growing exponentially for many generations, and where

all cellular component thus theoretically increases by the same constant factor per unit time (Ingraham *et al.*, 1983). The importance of balanced growth for standardization of bacterial cultures, as originally proposed by the Copenhagen School of Bacterial Growth Physiology (Schaechter, 2006), has thus in the present study been validated. The conclusions strongly suggest that physiological experiments follow the guidelines of the Copenhagen school.

## Acknowledgements

This work was supported by the Danish Council for Strategic Research. The technical assistance from Marzanna Pulka-Amin is highly appreciated. We thank Inka Sastalla for the kind gift of plasmid pSW4-GFPopt. We thank Flemming Hansen for fluorescent proteins guidance.

## 3.6 References

- Albano, C. R., Randers-Eichhorn, L., Bentley, W. E. & Rao, G. (1998).** Green fluorescent protein as a real time quantitative reporter of heterologous protein production. *Biotechnol Prog* **14**, 351–4.
- Bertani, G. (1951).** Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* **62**, 293–300.
- Boylan, R. J., Mendelson, N. H., Brooks, D. & Young, F. E. (1972).** Regulation of the bacterial cell wall: analysis of a mutant of *Bacillus subtilis* defective in biosynthesis of teichoic acid. *J Bacteriol* **110**, 281–290.
- Burkholder, P. R. & Giles, N. H. (1947).** Induced biochemical mutation in *Bacillus subtilis*. *Am J Bot* **34**, 345–348.
- Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G. & Ward, W. W. (1993).** Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* **32**, 1212–8.
- Condon, C., Squires, C. & Squires, C. L. (1995).** Control of rRNA transcription in *Escherichia coli*. *Microbiol Rev* **59**, 623–45.
- Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996).** FACS-optimized Mutants of Green Fluorescent Protein (GFP). *Gene* **173**, 33–38.
- Van Dijl, J. M. & Hecker, M. (2013).** *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact* **12**.
- Ecker, R. E. & Schaechter, M. (1963).** Ribosome content and the rate of growth of *Salmonella Typhimurium*. *Biochim Biophys Acta* **76**, 275–279.
- Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. (2002).** Stochastic gene expression in a single cell. *Science* **297**, 1183–1186.
- Gourse, R. L., Gaal, T., Bartlett, M. S. J., Appleman, A. & Ross, W. (1996).** rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol* **50**, 645–77.

- Ingraham, J. L., Maaløe, O. & Neidhardt, F. C. (1983).** *Growth of the Bacterial Cell*. Sunderland: Sinauer Associates Inc.
- Ishihama, A. (1997).** Adaptation of gene expression in stationary phase bacteria. *Curr Opin Genet Dev* **7**, 582–588.
- Ishihama, A. (1999).** Modulation of the nucleoid, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival. *Genes to Cells* **4**, 135–143.
- Kjeldgaard, N. O., Maaloe, O. & Schaechter, M. (1958).** The transition between different physiological states during balanced growth of *Salmonella typhimurium*. *J Gen Microbiol* **19**, 607–616.
- Kolter, R., Siegele, D. a & Tormo, A. (1993).** The stationary phase of the bacterial life cycle. *Annu Rev Microbiol* **47**, 855–874.
- March, J. C., Rao, G. & Bentley, W. E. (2003).** Biotechnological applications of green fluorescent protein. *Appl Microbiol Biotechnol* **62**, 303–315.
- Martínková, L., Machek, F., Ujocová, E., Kolín, F. & Zajíček, J. (1991).** Effect of age, amount of inoculum and inoculation medium composition on lactic acid production from glucose by *Lactobacillus casei* subsp. *rhamnosus*. *Folia Microbiol* **36**, 246–248.
- McAdams, H. H. & Arkin, a. (1997).** Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A* **94**, 814–819.
- Miura, a, Krueger, J. H., Itoh, S., de Boer, H. a & Nomura, M. (1981).** Growth-rate-dependent regulation of ribosome synthesis in *E. coli*: expression of the *lacZ* and *galK* genes fused to ribosomal promoters. *Cell* **25**, 773–782.
- Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. (1990).** *Physiology of the bacterial cell: a molecular approach*. Sunderland, Massachusetts: Sinauer Associates Inc.
- Neidhardt, F. C. & Magasanik, B. (1960).** Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim Biophys Acta* **42**, 99–116.
- Neves, a a, Vieira, L. M. & Menezes, J. C. (2001).** Effects of preculture variability on clavulanic acid fermentation. *Biotechnol Bioeng* **72**, 628–633.
- Roszak, D. B. & Colwell, R. R. (1987).** Survival strategies of bacteria in the natural environment. *Microbiol Rev* **51**, 365–379.
- Samarrai, W., Liu, D. X., White, A.-M., Studamire, B., Edelstein, J., Srivastava, A., Widom, R. L. & Rudner, R. (2011).** Differential responses of *Bacillus subtilis* rRNA promoters to nutritional stress. *J Bacteriol* **193**, 723–733.
- Sambrook, J. & Russell, D. W. (2001).** *Molecular Cloning A Laboratory Manual*.
- Sargent, M. G. (1975).** Control of Cell Length in *Bacillus subtilis*. *J Bacteriol* **123**, 7–19.
- Sastalla, I., Chim, K., Cheung, G. Y. C., Pomerantsev, A. P. & Leppla, S. H. (2009).** Codon-optimized fluorescent proteins designed for expression in low-GC gram-positive bacteria. *Appl Environ Microbiol* **75**, 2099–2110.

- Saxild, H. H. & Nygaard, P. (1987).** Genetic and physiological characterization of *Bacillus subtilis* mutants resistant to purine analogs. *J Bacteriol* **169**, 2977–2983.
- Schaechter, M., Maaloe, O. & Kjeldgaard, N. (1958).** Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *Microbiology* **19**, 592–606.
- Schaechter, M. (2006).** From growth physiology to systems biology. *Int Microbiol* **9**, 157–161.
- Scholz, O., Thiel, A., Hillen, W. & Niederweis, M. (2000).** Quantitative analysis of gene expression with an improved green fluorescent protein. *Eur J Biochem* **267**, 1565–1570.
- Sharpe, M. E., Hauser, P. M., Sharpe, R. G. & Errington, J. (1998).** *Bacillus subtilis* cell cycle as studied by fluorescence microscopy: Constancy of cell length at initiation of DNA replication and evidence for active nucleoid partitioning. *J Bacteriol* **180**, 547–555.
- Silva-Rocha, R. & de Lorenzo, V. (2012).** A GFP-*lacZ* bicistronic reporter system for promoter analysis in environmental gram-negative bacteria. *PLoS One* **7**, e34675.
- Slatkin, M. (1974).** Hedging one's evolutionary bets. *Nature* **250**, 704–705.
- Spizizen, J. (1958).** Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc Natl Acad Sci* **44**, 1072–1078.
- Sternberg, C., Christensen, B. B., Johansen, T., Toftgaard Nielsen, A., Andersen, J. B., Givskov, M. & Molin, S. (1999).** Distribution of bacterial growth activity in flow-chamber biofilms. *Appl Environ Microbiol* **65**, 4108–4117.
- Thyregod, C., Pedersen, A. E., Kilstrup, M. & Martinussen, J. (2015).** Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry. *in preparation*.
- Westers, L., Westers, H. & Quax, W. J. (2004).** *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta* **1694**, 299–310.

## 4 Heterogeneity in populations of *L. lactis*

### 4.1 Introduction

The oval cocci shaped lactic acid bacteria *L. lactis* is massively utilized as starters in the dairy industry (Holzapfel & Wood, 2014) and as an industrial production organism for specific products like lactic acid (John *et al.*, 2007), bacteriocins (Rodríguez *et al.*, 2003), as probiotics (Miyoshi *et al.*, 2002), and as vaccine delivery vehicle (Wyszyńska *et al.*, 2015).

The industrial relevance and the fact that *L. lactis* is a model organism for low GC content Gram-positive microorganisms the study of population heterogeneity, in this particular organism, is not only interesting but also highly relevant.

In order to study phenotypic variation in *L. lactis* using green fluorescent protein some issues regarding the reporter strain and the experimental setup needed to be considered. As mention in the introduction oxygen is required for maturation of the chromophore. The dependence of GFP's oxidation rate on oxygen concentration is not well elucidated (Tsien, 1998), thus it is difficult to determine at how low an oxygen concentration the activation of the GFP chromophore would be affected. *L. lactis* is a facultative anaerobe, and is propagated without vigorously shaking of the growth medium, but in the presence of atmospheric oxygen. The culture is stirred with a small magnet to preventing precipitation. It is not known whether the maturation of the chromophore would be affected by these cultivation conditions. In 2001 Hansen *et al.* showed that fluorescence was emitted from a *Streptococcus gordonii* GFP producing strain cultivated under anaerobic conditions, if a reducing agent was not added in the medium. Without a reducing agent the medium still contained 0.1 p.p.m. dissolved oxygen, which evidently was enough for maturation of the GFP chromophore. If the reducing agent L-cysteine was added to the medium the dissolved oxygen concentration was lowered to less than 0.025 p.p.m (lower detection limit). Under this condition the GFP expressing *S. gordonii* was not fluorescent (Hansen *et al.* 2001).

In the experimental setup used to produce the results presented in this section *L. lactis* was cultivated in 100 ml flasks with 100 ml medium and a small magnetic stirrer rotating to prevent precipitation of the cells. The dissolved oxygen was not measured, but fluorescent microscopy images revealed that all cells were fluorescent during balanced growth and in the consecutive stationary phase (data not shown).

The pH of the media is important to consider when expressing GFP in lactic acid bacteria, as the fluorescence of GFP decreases in acidic environments. At which pH the fluorescence gets affected depends on the variant of GFP; wild type GFP (wtGFP) is stable at a pH above 6, the fluorescence intensity from other variants already decreases when pH gets below 7 (Patterson *et al.*, 1997). In the article from 2001 Hansen *et al.* showed that the fluorescence intensity of a GFP expressing *S. gordonii* cultivated in anaerobic condition (without reducing agent in the medium), is affected as the pH approaches 6.5.

Lactic acid bacteria acidify the medium during growth, and as the external pH (the pH of the medium) decreases the internal pH of *L. lactis* also decreases. As a rule of thumb, the internal pH is one pH unit higher than the external pH, below pH 6 (Hutkins & Nannen, 1993; Mercade *et al.*, 2000; Nannen & Hutkins, 1991).

## **4.2 Materials and methods – *L. lactis***

### **4.2.1 Media**

*L. lactis* was propagated in GM17 (complex medium) or in the chemically defined SA-medium (Jensen & Hammer, 1993). SAL medium, SA medium where Na-acetated is replaced with liporic acid [2 mg L<sup>-1</sup>], was supplemented with 5 µg/ml erythromycin and the final concentration of morpholinepropanesulfonic acid (MOPS) was increased to 200 mM (instead of 40 mM as the protocol describes). The concentration of MOPS in the medium was increased to obtain a neutral pH during the experiments. A concentration of 200 mM MOPS in SA medium would have an affect the growth rate of *L. lactis* MG1363, but balanced growth could still be attained (Ole Michelsen personal communication, and (Jensen & Hammer, 1993)).

The pH of the SALM medium (SAL medium with increased concentration of MOPS) was adjusted to 7.40 prior to filter sterilization. The carbon source was added to a final concentration of 0.5 %. Cultivation of *L. lactis* was done in flasks with a small magnetic stirrer, preventing cell precipitation and ensuring some aeration. SALM were supplemented with glucose, maltose or with glucose and nucleosides (uridine 20 µg/ml, guanosine 20 µg/ml, thymidine 20 µg/ml, hypoxanthine 20 µg/ml and cytidine 20 µg/ml).

The SALM will be referred to as chemically defined medium in connection with cultivation of *L. lactis*.

Cultivation of *L. lactis* MG1363 was done at 30°C. Samples were, immediately after sampling, placed on ice and analyzed within minutes.

### **4.2.2 Experimental setup for making a balanced growing *L. lactis* culture**

A single colony was inoculating in prewarmed chemically defined medium. The single colony was propagated in 6-8 hours at 30 °C. After 6-8 hours of growth the pre-pre-culture was used as inoculum for a dilution series which was propagated overnight. The inoculum used to make a culture in balanced growth was an overnight culture from the dilution series, with an OD<sub>600</sub> between 0.2 and 0.4.

### **4.2.3 Fluorescence microscopy, spectrofluorophotometry and flow cytometry**

The spectrofluorophotometer (RF-5301PC Shimadzu) was used for measurements of fluorescence intensity on a population level. The emission and excitation slit were set to 5. The excitation wavelength was set at 488 nm and the emission spectrum was measured. Emission intensity was determined at 508 nm. SALM medium had a considerable auto fluorescent and was thus removed before spectrofluorimetric analysis of the *L. lactis* samples. The culture samples were spun down at 4000 rpm for 4 minutes in a refrigerated centrifuge.



The supernatant was discarded and the cells were resuspended in ice-cold PBS solution (0.137M NaCl, 0.003M KCl, 0.012M NaHPO<sub>4</sub> and 0.002M KH<sub>2</sub>PO<sub>4</sub> pH 7.2) to an OD<sub>600</sub> between 0.2 and 0.4 before spectrofluorophotometer measurements.

The control strain AE071 exhibited a measurable, but close to the detection limited of the spectrofluorophotometer, autofluorescence. Due to the fact that the autofluorescence of control strain was close to (and often below) the detection limited of the apparatus, it was neglected.

Fluorescence microscopy was done using a Zeiss Axioplan fluorescence microscope. GFP fluorescence was analyzed with the use of Semrock GFP-3035C BrightLine single-band filter set. Fluorescence microscopy was used to isolate transformants.

A FACSCalibur (BD Biosciences), with a 15 mW, 488 nm, air-cooled argon-ion laser was used for single cell analysis. Fluorescence emission levels were measured using a band pass filter at 530/30 nm (FITC). The settings used for analysis of heterogeneity in *L. lactis* cultures inoculated with either an exponentially growing or a stationary phase pre-culture were: FSC E01, SSC 375 V, and FL1 906 V. A side scatter threshold was applied to gate out much of the noise (at channel 144). The settings FSC E02, SSC 505 V, and FL1 860 V were used for analysis of *L. lactis* when propagated in chemically defined medium supplemented with glucose, maltose or glucose plus nucleoside and for analysis of the experiment where the *L. lactis* reporter strain was propagated in either complex medium or chemically defined medium. All parameters were logarithmically amplified and a side scatter threshold was applied to gate out much of the noise (at channel 144).

During acquisition of *L. lactis* cultures, the fluorescent events were gated (all events were saved), the software counted 100,000 fluorescent events before ending acquisition.

Histogram gates were applied in the post-processing of the flow cytometer data according to fluorescence. Histogram gating was done using Flowing Software (v2.5.1, Turku Centre for Biotechnology, Turku, Finland), before statistical analysis was done in JMP®. FlowJo® was used for graphical display of the data (contour plots).

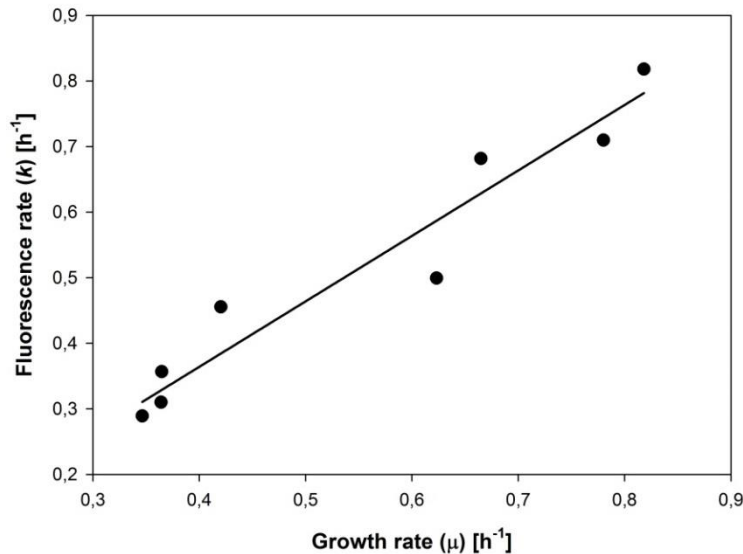
## 4.3 Results and discussion

### 4.3.1 GFP expression in the fluorescent *L. lactis* reporter strain was growth rate regulated

To determine if the GFP expression in the *L. lactis* reporter strain AE072 was dependent on growth rate the reporter strain was propagated at different growth rates and the fluorescence was measured.

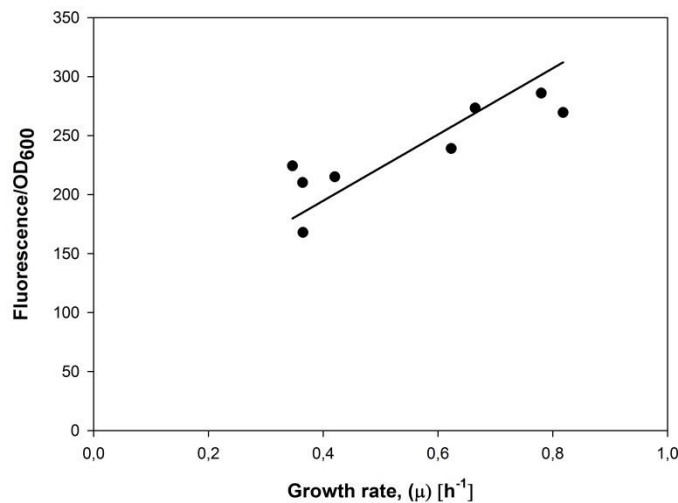
The spectrofluorimetric measurements of fluorescence were complicated by the high autofluorescence of the chemically defined medium and the cells consequently had to be resuspended in PBS solution before analysis. The chemically defined medium contained riboflavin which could be one of the components contributing to the large autofluorescence observed for this medium (Billinton & Knight, 2001). Because the cells had to be resuspended in PBS solution before analysis, the fluorescence at a specific cell density ( $FL_{medium}$ ) was determined by dividing the fluorescence measured for the cells resuspended in PBS ( $FL_{PBS}$ ), with the PBS sample cell density ( $OD_{600, PBS}$ ) and then multiply with the optical density of the cell culture ( $OD_{600, medium}$ ) at the time the sample was taken. 
$$\left( \frac{FL_{PBS}}{OD_{600, PBS}} \cdot OD_{600, medium} = FL_{medium} \right)$$

The calculated fluorescence ( $FL_{medium}$ ) was plotted as a function of time to determine the specific fluorescence rate ( $k, h^{-1}$ ), see Figure 8



**Figure 8** Proportionality between growth rate and specific fluorescence rate was determined on a population level. The plot shows the rates at which the fluorescence intensity increases (specific fluorescence rate,  $k, [h^{-1}]$ , vertical axis), at various growth rates ( $\mu, [h^{-1}]$ , horizontal axis). Growth rates were determined with optical density measurements. The fluorescence was measured with a RF-5301PC Shimadzu spectrofluorophotometer. The chemically defined SALM medium was supplemented with galactose, maltose, glucose or glucose plus nucleoside in order to obtain different growth rates. Regression line:  $y = 0.9807x - 0.022$   $R^2=0.93$

It was observed that the specific fluorescence rate  $k$  ( $\frac{dFL}{dt} = k FL$ ) was proportional to the growth rate, indicating the biomass and fluorescence increase at the same rate (Figure 8). In *L. lactis* are ribosomal promoters growth rate regulated (Beresford & Condon, 1993) and the expression of GFP was thus anticipated to be regulated by growth rate as GFP was transcriptional regulated by the 16S ribosomal promoter  $P_{rmC}$ . The specific fluorescence (fluorescence per OD) was plotted against the growth rate (Figure 9). The specific fluorescence correlated with the growth rate, and it was concluded that GFP expression in the *L. lactis* reporter strain AE072 was growth rate regulated on a population level.



**Figure 9 Correlation between growth rate and specific fluorescence. Regression line  $175.79x + 139.5$   $R^2 = 0.7653$ . The data used in this plot originated slope of the linear dependency between fluorescence and optical density plotted against the growth rate of the culture.**

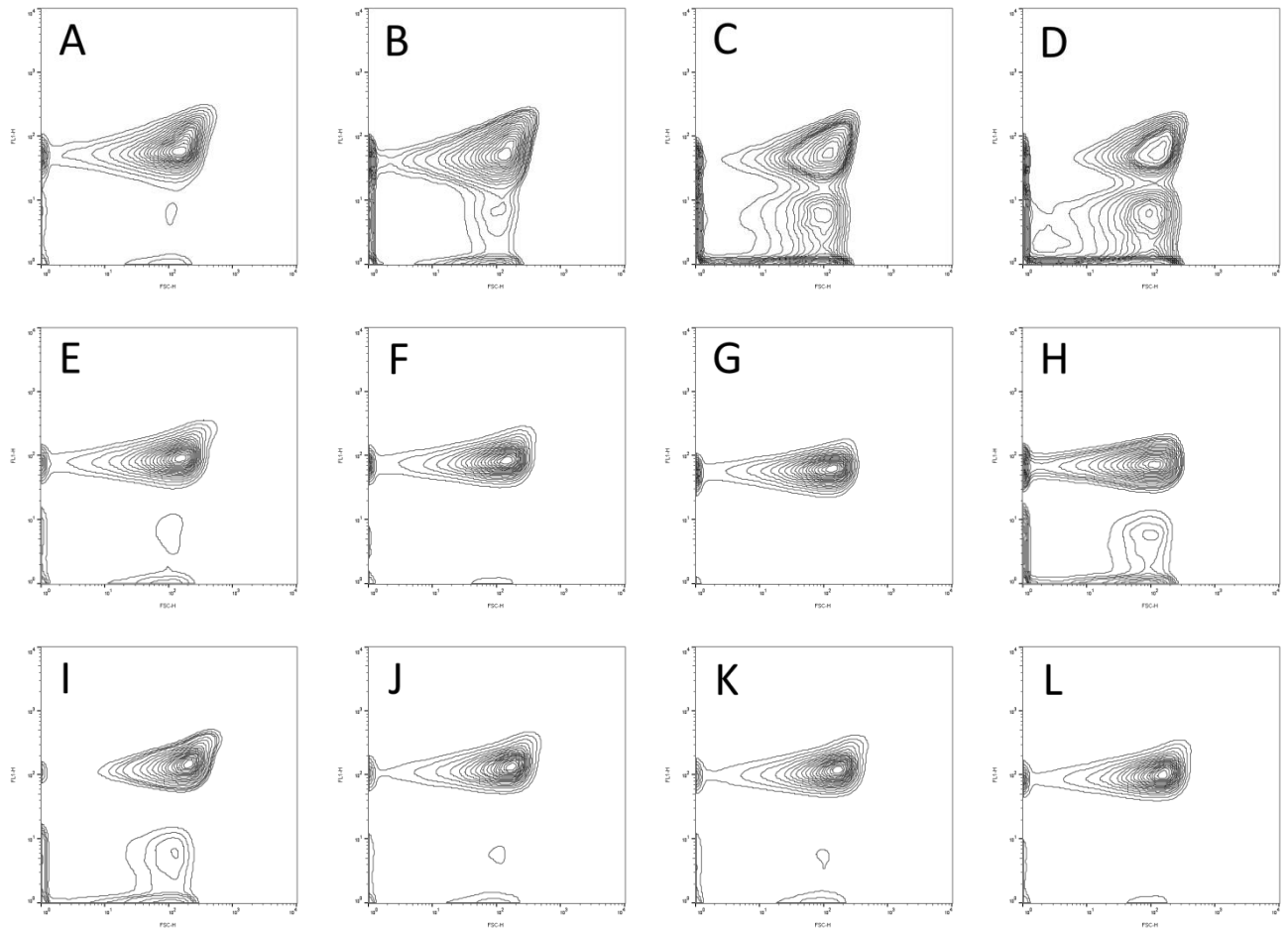
The pH of the chemically defined medium, with the increased concentration of MOPS, was determined after several of the experiments. The cultures had entered stationary phase before pH was determined. The pH was consistently above 6.8, indicating that the pH was kept close to or above 7 at all-time during cultivation in exponential phase. Besides increasing the buffer concentration to a maximum, it would be difficult to take further actions towards maintaining a neutral pH. Based on the fact that the pH of the medium never gets below 6.8 during cultivation and that the samples were resuspended in PBS solution (pH 7.2) before analysis, it was concluded that the effect on the fluorescence due to decrease in pH was minimal.

### 4.3.2 Population heterogeneity at different growth rates

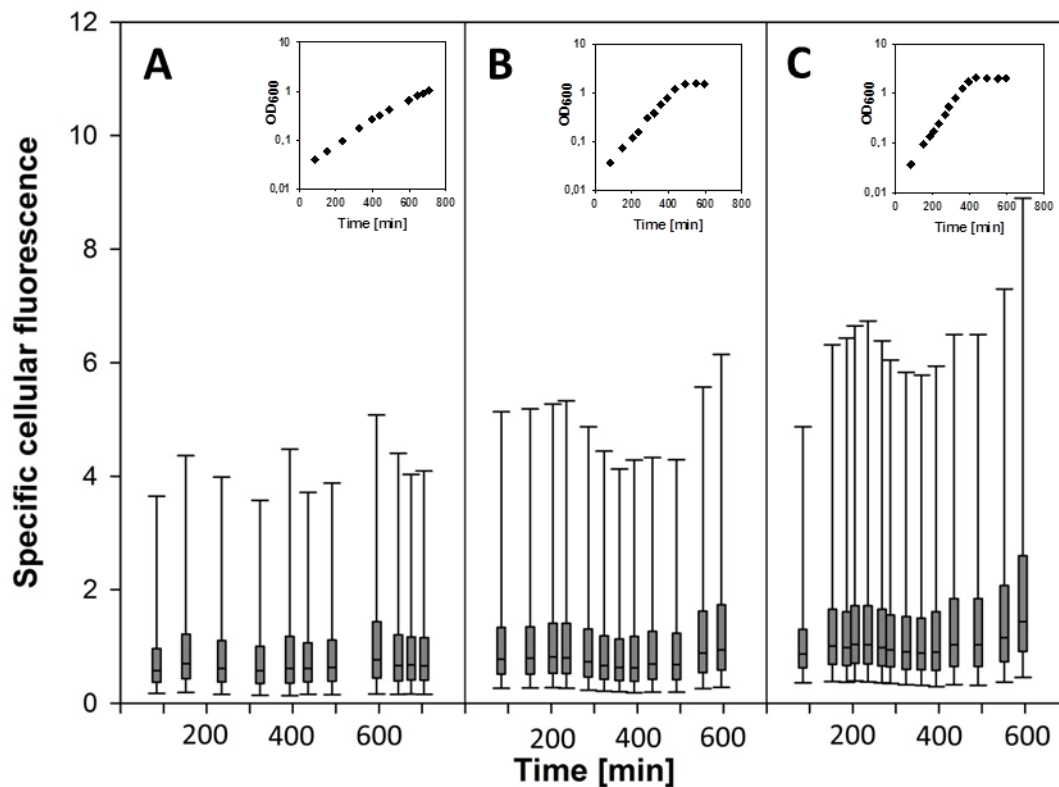
The green fluorescent reporter strain with growth rate regulated expression of GFP was cultivated as balanced growing cultures at three different growth rates and analyzed by flow cytometry, to investigate cell-to-cell variability with focus on single cell growth rate.

The flow cytometry data presented in contour plots (Figure 10) shows the population profiles for AE072 when cultivated at 133 minutes, 70 minutes and 54 minutes doubling. The shape of the population was noticeable different when the cells were propagated in medium supplemented with maltose (doubling time 133 minutes). It was observed that the size of *L. lactis* cells, with the same fluorescence intensity, varied appreciable at all three growth rates – this tendency will be discussed further in section 4.3.3.

Specific cellular fluorescence was computed and the result is shown as box-plots in Figure 11. A log-normal distribution and a normal distribution were fitted to the data (fitted in JMP®, data not shown). The distribution was closest to a log-normal distribution, which can also be visualized from the shape of the box-plots, as the median is not located in the center of the box. From the data presented in the supplementary materials (section 11.3.1) it was observed that the average cell size was not affected markedly by the growth rate, the cells could have a tendency to be slightly bigger when they are dividing faster, or more duplets or longer chains of cells was simply present at faster growth.



**Figure 10** Contour plots showing the distribution of size (forward scatter, FSC-H, horizontal axis) and fluorescence (FL1-H, vertical axis) attained for the *L. lactis* reporter strain AE072 when propagated as balanced cultures at different growth rates. (A – D) AE072 cultivated in chemically defined medium with maltose, doubling time 133 minutes. The samples presented in the contour plots were taken 0.8 generation (84 minutes), 2.9 generation (324 minutes), 4.1 generation (491 minutes) and 5.1 generations (645 minutes) after inoculation, respectively. All samples were taken during exponential growth. (E – H) AE072 cultivated in chemically defined medium with glucose, doubling time 70 minutes. The samples were taken 1.2 generations (84 minutes), 3.3 generation (235 minutes), 5.1 generations (358 minutes) after inoculation with the exponentially growing pre-culture. The sample presented in plot H was taken in stationary phase (552 minutes after inoculation). (I – L) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes. The samples were taken 1.2, 3.4 and 5.1 generations after inoculation (84 minutes, 205 minutes and 287 minutes after inoculation). The sample shown in contour plot L was taken when the culture had entered stationary phase.



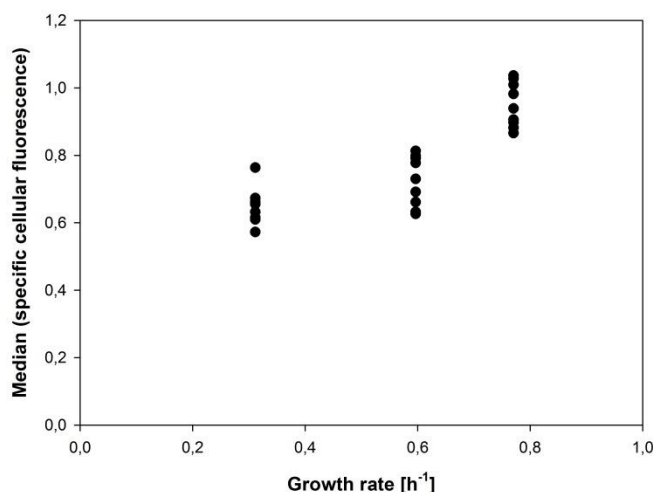
**Figure 11** Box-plot showing specific cellular fluorescence for the *L. lactis* reporter strain AE072 when cultivated at three different growth rates. (A) AE072 cultivated in chemically defined medium supplemented with maltose, doubling time 133 minutes. (B) AE072 cultivated in chemically defined medium supplemented with glucose, doubling time 70 minutes. (C) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes. Growth curves are shown above the box-plot. Each diamond symbol in the growth curves can be linked to a box-plot in A, B and C, respectively.

Fluctuation between the samples can be observed for the coefficient of variation for the slow growing culture, both for specific cellular fluorescence and the forward scatter per cell (Figure 54, bottom left, A). This tendency indicates a higher level of heterogeneity in the slow growing culture.

When a population of *L. lactis* GFP reporter cells entered stationary phase the specific cellular fluorescence standard deviation, mean and median values, and interquartile range increased. This observed increase in cell-to-cell variability could be expected as the cells encounter many morphological and physiological changes when entering stationary phase, factors which increase the phenotypic noise of the cell. But tendencies observed in stationary phase will be affected by several parameters: GFP is quite stable and due to the transcriptional regulation in the reporter strain will the fluorescence emitted from stationary phase cell originate from GFP, which had been expressed during exponential growth. To disturb the picture even more

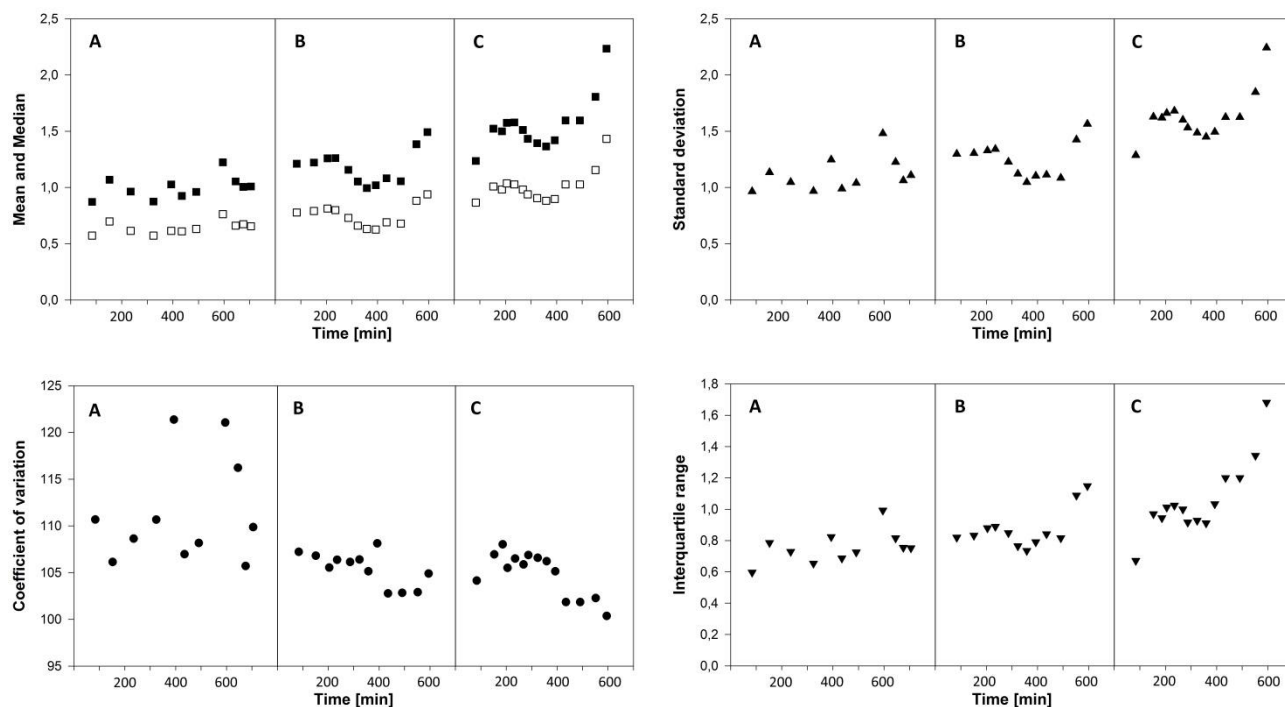
will GFP expressed in late-exponential phase first become fluorescent later, due to maturation delay and thus first be observable in stationary phase. These two effects must then be combined with a decrease in cell size, which was observed from forward scatter measurements (see data for forward scatter per cell in supplementary materials, section 11.3.1). Using the descriptive statistics presented in Figure 13 the skewness can be observed, as the mean and median values differ. Analysis of the distribution were done in JMP® which, for the chemically defined medium presented in this section, showed to fit a log-normal distribution better than a normal distribution (data not shown). The standard deviation (Figure 13, top right) indicate that cell-to-cell variation was largest in the fastest growing culture. But as the data tends to be log-normal distributed it would be more optimal to use the coefficient of variation (Figure 13, bottom left), to define the level of heterogeneity, as the standard deviation of a log-normal distribution will increase as the mean value increases.

The *L. lactis* reporter strain was propagated at three different growth rates to demonstrate that the specific cellular fluorescence could be used as a measure for a single cell growth rate. The mean and median values determined for the samples are shown in Figure 13(top left). The specific cellular fluorescence sample median values, for the samples taken during exponential phase, are in Figure 12 plotted against the three growth rates. The median values for the fastest growing culture ( $\mu$ :  $0.77 \text{ h}^{-1}$ ) are larger than the median values obtained for the culture propagated in medium supplemented with glucose ( $\mu$ :  $0.60 \text{ h}^{-1}$ ). The differences between the median values for the cultures propagated in medium supplemented with glucose and cultures propagated in medium supplemented with maltose ( $\mu$ :  $0.31 \text{ h}^{-1}$ ) are almost absent.



**Figure 12** Correlation between growth rate and specific cellular fluorescence. The median values determined for each sample is plotted against the growth rate of the culture. Growth rate for the culture cultivated in medium supplemented with maltose  $0,31 \text{ h}^{-1}$ . Growth rate for the culture propagated in medium supplemented with glucose:  $0,60 \text{ h}^{-1}$  and the growth rate for the culture propagated in medium supplemented with glucose and nucleosides:  $0,77 \text{ h}^{-1}$ .

Looking at the results obtained for AE072 cultivated at different growth rates it can be concluded that a certain level of heterogeneity was unavoidable. Whether the cell-to-cell variability considering single cell growth rate was highest when the cells were dividing at a fast rate or at a slower rate will require more investigation. Perhaps it would be difficult to resolve due to the population profiles obtained for the *L. lactis* GFP reporter strain, when cultivated in chemically defined medium (see i.a. the contour plots Figure 10). This will be discussed further in the next section.



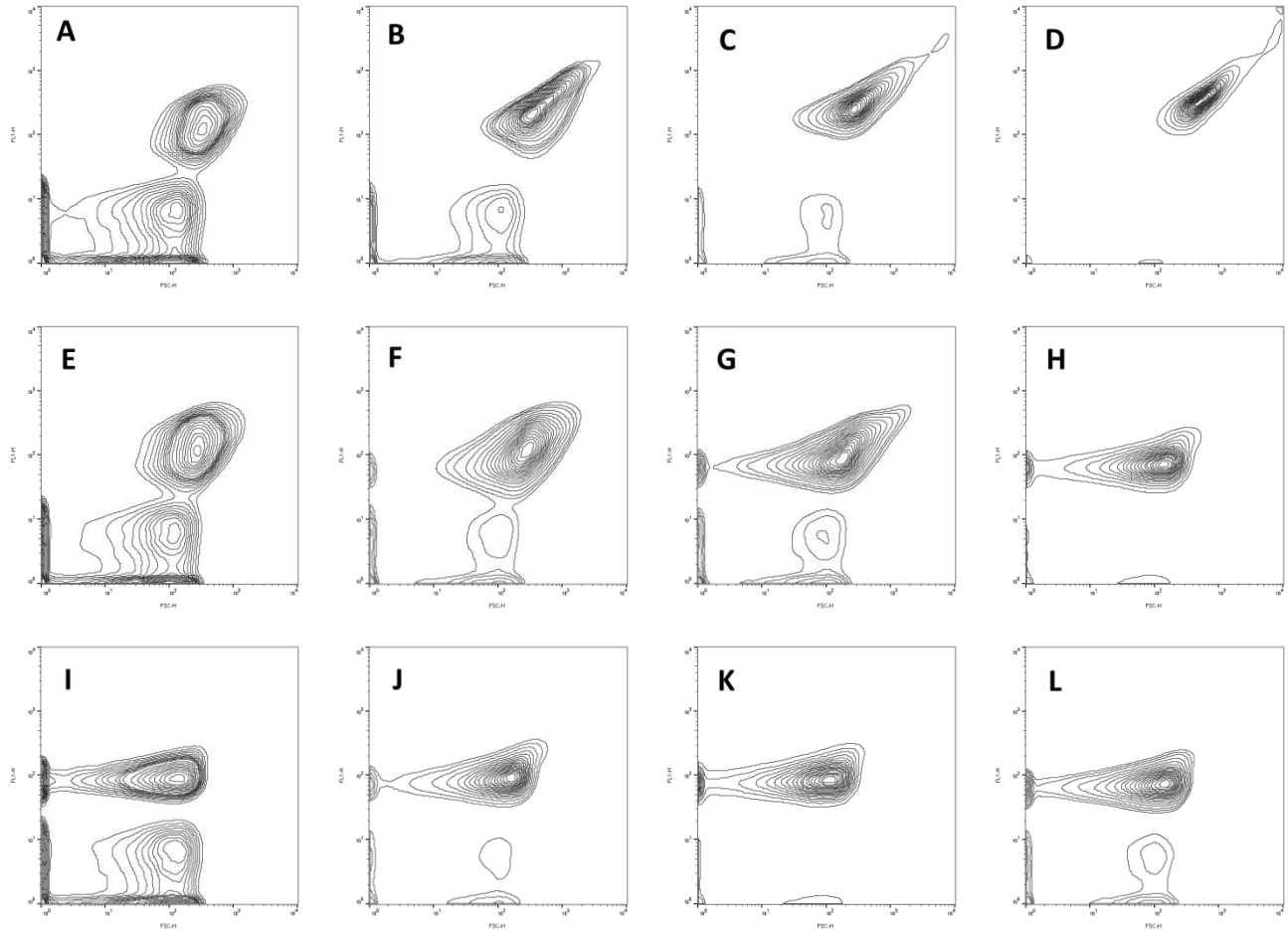
**Figure 13** Descriptive statistics for the specific cellular fluorescence measured for the *L. lactis* GFP reporter strain cultivated in chemically defined medium supplemented with (A) maltose, doubling time 133 minutes, (B) glucose, doubling time 70 minutes and (C) glucose plus nucleosides doubling time 54 minutes. Growth curves can be found in Figure 11. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down.

### 4.3.3 Complex medium changes the shape of the population

Cell-to-cell variability was, as shown in the section above, present in the balanced growing *L. lactis* cultures propagated at different growth rates. In *B. subtilis* it was observed that the level of population heterogeneity clearly was higher, if a stationary phase pre-culture was used as inoculum. To elucidate if this also was the case for *L. lactis* the following experiment was conducted: a single colony from a complex medium (GM17) agar plate was inoculated into complex medium. This complex medium pre-culture was left to grow overnight. Complex medium and chemically defined medium (SALM medium supplemented with glucose) was inoculated with the stationary phase complex medium pre-culture. The stationary phase overnight culture was



diluted in the ratio 1:62.5. For comparison a balanced growing culture was made in chemically defined medium as described in section 4.2.2.



**Figure 14** Flow cytometry data for the *L. lactis* reporter strain AE072 propagated at three different experimental conditions shown as contour plots. Forward scatter per cell (horizontal axis) and fluorescence per cell (vertical axis). (A - D) *L. lactis* reporter strain AE072 cultivated in complex medium inoculated with a complex medium stationary phase pre-culture, doubling time in exponential phase: 46 minutes. The samples presented as contour plots were taken 38 minutes, 116 minutes, 219 minutes and 265 minutes after inoculation. The growth curve can be seen in Figure 15A. (E - H) *L. lactis* reporter strain AE072 cultivated in glucose chemically defined medium inoculated with a complex medium stationary phase pre-culture, doubling time in exponential phase: 74 minutes. The samples presented as contour plots were taken 39 minutes, 93 minutes, 162 minutes and 410 minutes after inoculation, the growth curve can be seen in Figure 15B. (I - L) *L. lactis* reporter strain AE072 cultivated in glucose chemically defined medium inoculated with an exponentially growing glucose chemically defined medium pre-culture (balanced growth). The samples were taken 40 minutes, 142 minutes, 245 minutes and 376 minutes after inoculation. This culture had a 70 minutes doubling time during exponential growth. The growth curve can be seen in Figure 15C.

In Figure 14 it can be observed that complex medium (Figure 14A-D) supports a different population profile than chemically defined medium (Figure 14I-L). The stationary phase complex medium pre-culture used, for inoculum, resulted in comparable profiles in complex medium (Figure 14A) and in chemically defined medium (Figure 14E) 40 minutes after inoculation. The population profiles in Figure 14F and Figure 14G show how the complex medium pre-culture adapts to the new medium and finally in Figure 14H ends up with a profile resembling a profile observed for the balanced growing chemically defined medium culture shown in Figure 14I-L. From the population profile of the balanced growing culture (Figure 14I-L) it can be seen that the size of *L. lactis* cells, with the same fluorescence intensity, varied appreciable. Using the parameter specific cellular fluorescence to analyze variation in single cell growth rate should probably be done tentatively, or not at all, due to this proportion. The flow cytometry data was instead used for analysis of variation in cell size and variation in single cell fluorescence. Fluctuation in cell size will, as discuss previously, also affect the fluorescence cell-to-cell variability, but distribution in fluorescence per cell will still indicate the degree of population heterogeneity. The variation in specific cellular fluorescence is shown as box-plot in supplementary materials (section 11.3.2).

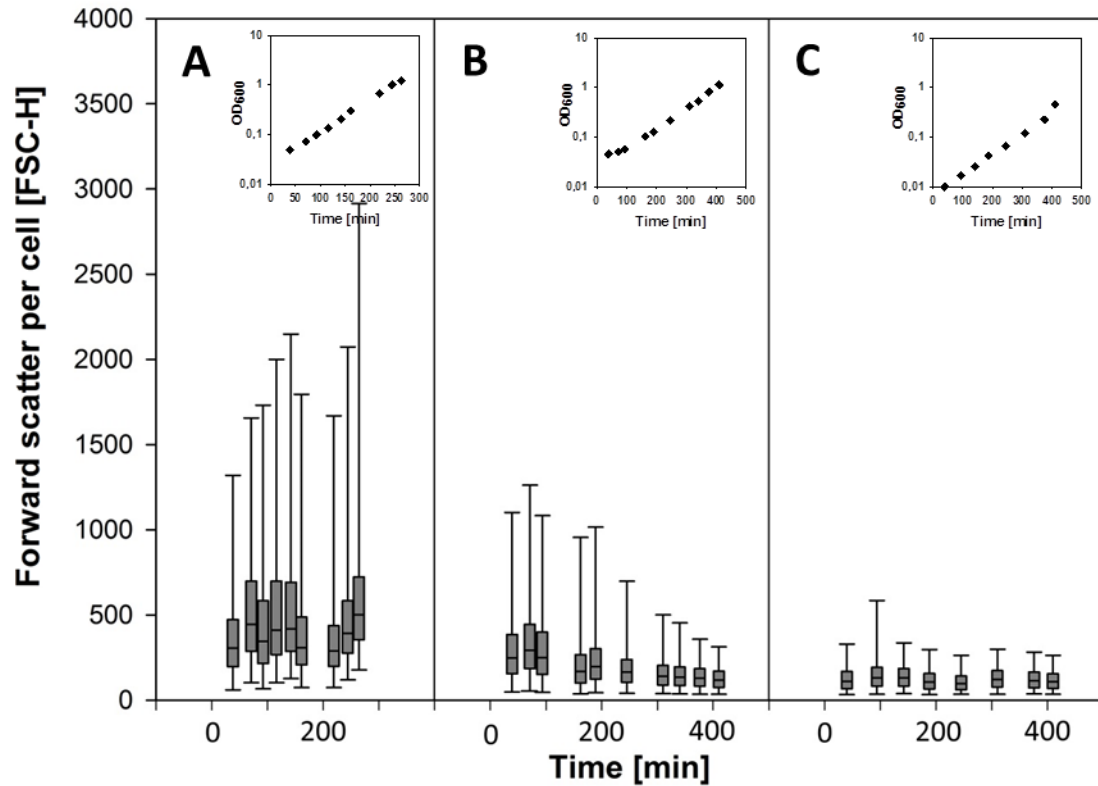
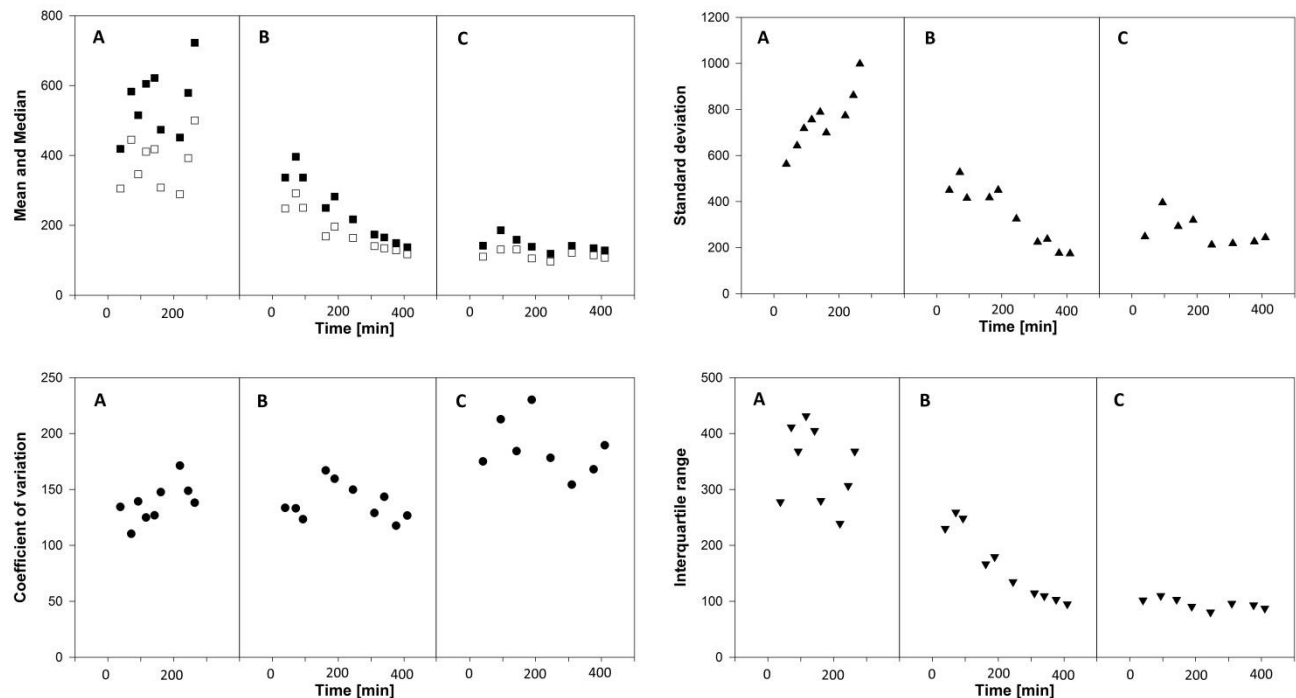


Figure 15 Box-plot showing the variation in forward scatter per cell for the *L. lactis* GFP reporter strain AE072 when cultivated at three different growth conditions. (A) A stationary phase complex medium pre-culture was inoculated in pre-warmed complex medium. (B) A stationary phase complex medium pre-culture was inoculated in pre-warmed chemically defined medium (SALM supplemented with glucose). (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose). Growth curves are shown above the box-plots, each diamond symbol in the growth curves can be linked to a box-plot.



**Figure 16** Descriptive statistics based on forward scatter per cell measured by flow cytometry. The *L. lactis* reporter strain AE072 when cultivated at three different growth conditions. (A) A stationary phase complex medium pre-culture was inoculated in pre-warmed complex medium. (B) A stationary phase complex medium pre-culture was inoculated in pre-warmed chemically defined medium (SALM supplemented with glucose). (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose). (Middle left) Mean: black squares, median white squares, (middle right) standard deviation: triangle up, (bottom left) coefficient of variation: black circle, (bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 15

The growth rate in the exponential phase was for the *L. lactis* reporter strain AE072 when propagated in complex medium (GM17) 46 minutes. The growth rate in chemically defined medium was 74 minutes, when inoculated with a stationary phase complex medium pre-culture and 70 minutes when propagated as a balanced growing culture. The *L. lactis* cells are considerable larger when propagated in GM17. Interestingly the cells in the chemically defined medium inoculated from the stationary phase GM17 overnight culture, continue to be larger, during the first 200 minutes of the cultivation, than the cells in the balanced growing culture (Figure 15, middle left).

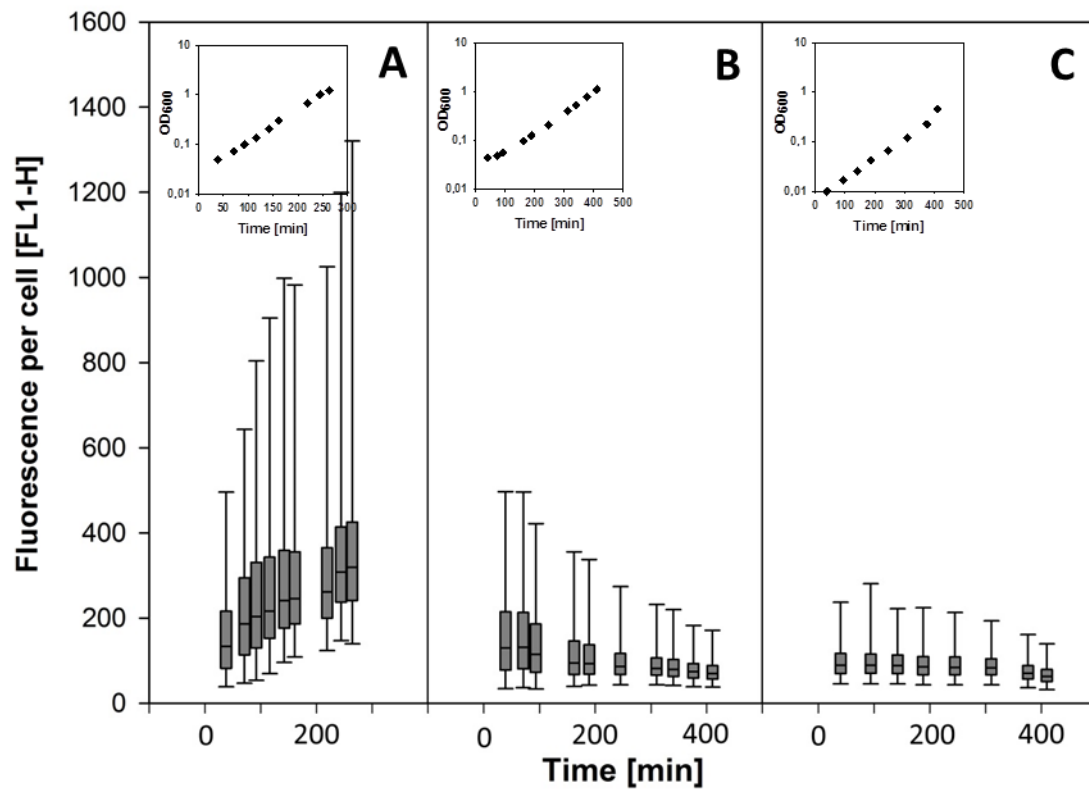
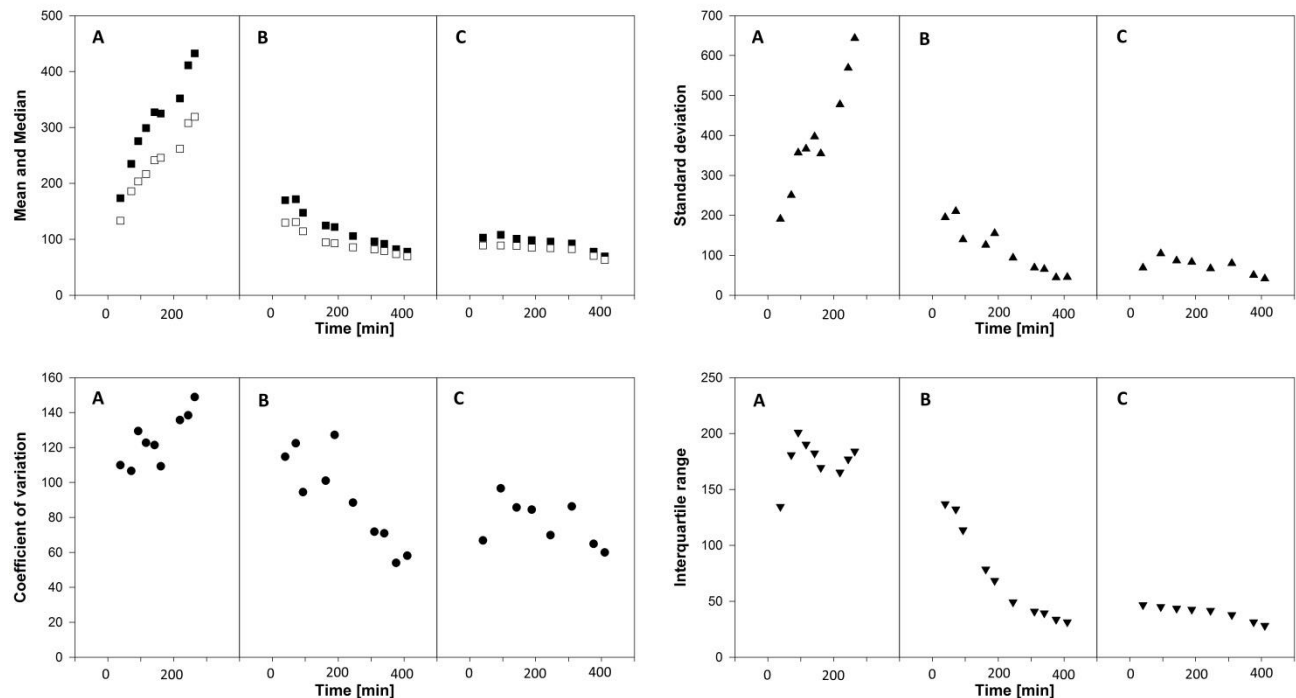


Figure 17 Box-plot showing the distribution of the fluorescence per cell determined for samples of *L. lactis* GFP reporter strain AE072 when cultivated at three different growth conditions. (A) A stationary phase complex medium pre-culture was inoculated in prewarmed complex medium. (B) A stationary phase complex medium pre-culture was inoculated in prewarmed chemically defined medium (SALM supplemented with glucose). (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose)



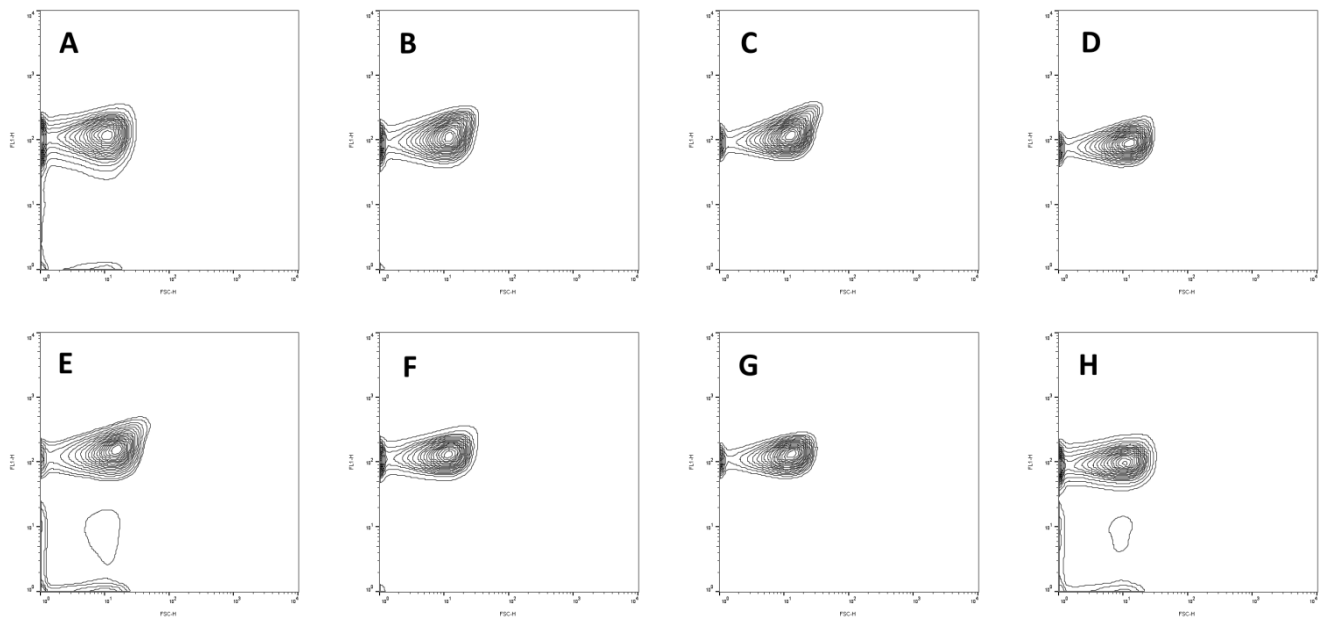
**Figure 18** Descriptive statistics based on fluorescence per cell observed for the *L. lactis* reporter strain AE072 when cultivated at three different growth conditions. (A) A stationary phase complex medium pre-culture was inoculated in prewarmed complex medium. (B) A stationary phase complex medium pre-culture was inoculated in prewarmed chemically defined medium (SALM supplemented with glucose). (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose). (Top left) Mean: black squares, median white squares, (Top right) standard deviation: triangle up, (bottom left) coefficient of variation: black circle, (bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 17.

From the results above it can also be concluded that cell-to-cell variability in cell size and cell fluorescence are markedly larger for the complex medium culture. The standard deviation and IQR determined in the respective samples all had increased values for the complex medium culture compared to the balanced growing culture.

Interestingly the chemically defined medium culture inoculated with the stationary phase complex medium pre-culture was, according to optical density measurements, in exponential growth approximately one hour after inoculation, but the cell size and fluorescence distribution continuously changed during the exponentially phase. The degree of heterogeneity in a sample taken from this culture, would depend on when the sample was taken during the exponential phase. If samples were taken from two independent balanced growing cultures, the population profiles of these would be similar and the degree of heterogeneity would likewise be similar.

#### 4.3.4 Population profiles complicate the analysis

From the results in the section above it was observed that the distribution of cell size and cellular fluorescence continuously change during exponential phase in the cultures inoculated with a stationary phase complex medium pre-culture. Results included in the manuscript in section 3 showed that the level of variation in single cell growth rates, in a population of *B. subtilis* cells was higher in cultures inoculated with a stationary phase minimal medium pre-culture than the level observed in a culture inoculated with an exponentially growing culture. To investigate if the state of the pre-culture also had an effect on the degree of heterogeneity in a *L. lactis* population two cultures, inoculated with a stationary phase inoculum and an exponentially growing inoculum respectively, were analyzed by flow cytometry. In Figure 19 are the population profiles shown as contour plots. The forward scatter was amplified less for this analysis than it had been during the flow cytometry analyses used to generate the results shown in the two preceding sections.



**Figure 19** Flow cytometry data for the GFP reporter strain AE072 shown in contour plots. Samples were taken from two chemically defined medium cultures, one inoculated with a stationary phase chemically defined medium pre-culture (A-D) and one inoculated with an exponentially growing chemically defined medium culture (E-H). Samples A-D were taken 77 minutes, 190 minutes, 349 minutes and 472 minutes after inoculation, respectively. Sample A was taken during the lag phase. Samples E-H were taken 5 minutes, 159 minutes, 313 minutes and 474 minutes after inoculation with the exponentially growing pre-culture. The doubling times for the two cultures were 68 and 66 minutes, respectively.

A general observation done analyzing flow cytometry data for *L. lactis* was that the diagonal shape of the population observed for *B. subtilis* and for the *L. lactis* culture propagated in complex medium (Figure 14 A-D) was not observed for the *L. lactis* strain when cultivated in chemically defined medium (SALM). The shape

observed when the *L. lactis* reporter strain was propagated in SALM medium was problematic to analyze using the computed parameter specific cellular fluorescence, as larger cell in general was not more fluorescence, a tendency which had been observed for *B. subtilis*. As the population profile was observed to be more diagonal for the *L. lactis* reporter strain when propagated in complex medium it could indicate that it is not an organism specific trait, it is an organism/medium combination trait. And this trait complicates the analysis of heterogeneity in single cell growth rates in chemically defined medium cultures of *L. lactis* AE072. The chemically defined medium population profiles shown in Figure 19, indicates that this organism/medium trait in combination with a non-optimal setting for the forward scatter signal makes the further analysis of these cultures complicated and the results would be characterized by sources of errors.

Speculation on what could affect the fluorescence in the cells the redox potential of *L. lactis* could be a potential source. Oxygenation of *L. lactis* cultures results in an altered redox state and greater NADH oxidase activity (Lopez de Felipe *et al.*, 1997, 1998). The NADH oxidase nox-1 reduce  $O_2$  to  $H_2O_2$  as NADH is reduced (Tachon *et al.*, 2011). Beside the increased  $H_2O_2$  formation will the oxygenation of the culture shift the sugar fermentation toward mixed fermentation (Lopez de Felipe *et al.*, 1997). In the reporter strain was the reactive oxygen species not only originating from NADH oxidase activity,  $H_2O_2$  is also formed in the ratio 1:1 in the formation of the GFP chromophore (Tsien, 1998; Zhang *et al.*, 2006b). *In vitro* experiments have shown that the relative fluorescence decreases as the  $H_2O_2$  concentration in the surrounding environment increases (Umakoshi *et al.*, 2011). And the fact that *L. lactis* is catalase negative and can thus not convert the formed  $H_2O_2$  to  $H_2O$  and  $O_2$  (Facklam & Elliot, 1995) and that the formation of the chromophore also contributes the  $H_2O_2$  formation, an intracellular concentration which could affect GFP fluorescence could be reached.

If the cells were not aerated enough, the maturation of GFP would, as mention in the introduction (section 4.1), be negatively affected. If the cells were aerated abundant the NADH oxidase activity would increase and the formation of reactive oxygen species (ROS) including  $H_2O_2$  would increase. Assuming that the cells were aerated enough to make all the GFP mature, the chromophore formation would contribute to formation of additional  $H_2O_2$ , perhaps making the concentration increase to a level which would have substantial effect on the cell, due to the toxicity of this compound.

GM17 medium supported a different population profile. The more diagonal profile observed in this medium could not alone have been caused by the higher growth rate, as a doubling time only 10 minutes lower was determined for the culture propagated in chemically defined medium supplemented with glucose and nucleosides. Thus other parameters must also have affected the GFP formation and/or the cell size during growth in the chemically defined medium.

To investigate this further the concentration of dissolved oxygen in the medium could be determined, the medium redox potential could be determined and the redox status of the cells could be analyzed by a redox sensing GFP (Hanson *et al.*, 2004). How the medium affect the population profile could be investigated by modifying the chemically defined medium by removing some of the non-essential amino acids and vitamins.



## 5 *S. cerevisiae* single cell analysis

### 5.1 Introduction

In the industry is *S. cerevisiae* used for production of various products - plant sesquiterpenes for perfume and pharmaceutical (Daviet & Schalk, 2010), heterologous proteins; human insulin, hepatitis virus vaccines, and human papilloma virus vaccines (Martínez *et al.*, 2012). Being an important production organism many lab scale fermentations and shake flask experiments are done with this yeast, for increase of knowledge and in attempt to increase productivity. Cell-to-cell variability have been demonstrated to for this organism, among others with the used of flow cytometry and fluorescent protein expressing reporter strains (Kacmar *et al.*, 2004; Keren *et al.*, 2015).

To investigate how the state of the pre-culture affects the degree of cell-to-cell variability in populations of the unicellular eukaryote *S. cerevisiae* the green fluorescent reporter strain FE440 was provided by request (Carlquist *et al.*, 2012). In this genetically manipulated yeast strain (FE440) the expression of GFP depends on the activity of the *rpl22a* promoter, a promoter which activity previously has been shown to be linearly correlated to growth rate (Brauer *et al.*, 2008; Carlquist *et al.*, 2012; Regenberg *et al.*, 2006). FE440 has previously been analyzed by flow cytometry (Carlquist *et al.*, 2012; Delvigne *et al.*, 2014).

In a paper from 2014 Delvigne *et al.* used FE440 in an on-line flow cytometry analysis of chemostat fermentations, with preliminary batch cultivation. The green fluorescent reporter strain was treated with PI before analysis. Delvigne *et al.* show that distribution of cellular fluorescence is bimodal during the preliminary batch fermentation and that the PI uptake (cells with damages membranes) was, during the first hours of the batch fermentation, relative high for thereafter to decrease to a level close to the detection limited (Delvigne *et al.*, 2014).

It had previously been observed that the GFP level of reporter strain FE440 changed during different growth phases of a batch fermentation; exponential growth on glucose, diauxic shift and exponential growth on ethanol, confirming that the growth reporter could be used to distinguish cells in different propagation modes (Carlquist *et al.*, 2012).

### 5.2 Materials and Methods

#### 5.2.1 Strains

*S. cerevisiae* strain FE522 (CEN.PK 113-5D with *ura3-52::URA3* chromosomal integration) and GFP reporter strain FE440 (CEN.PK 113-5D with *ura3-52::URA3-PRPL22A-yEGFP-TCYC1* chromosomal integration (Carlquist *et al.*, 2012)

### 5.2.2 Medium

*S. cerevisiae* was propagated in CBS (Centraalbureau voor Schimmelcultures) medium. For 1 liter of CBS chemically defined medium for yeast cultivation the following components were mixed: 7.5 g  $(\text{NH}_4)_2\text{SO}_4$ , 14.4 g  $\text{KH}_2\text{PO}_4$ , 0.5 g  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 50  $\mu\text{l}$  Sigma 204 antifoam, 2 mL Trace metals, 2-(N-morpholino)ethanesulfonic acid (MES) was added to a concentration of 100 mM. pH was adjusted to 6.5 before autoclavation. 1 mL of vitamins was added and carbon source was added to 2 %.

Solution of trace metals (1000 ml): 3.0 g  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 4.5 g  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ , 4.5 g  $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$ , 0.84 g  $\text{MnCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.3 g  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ , 0.4 g  $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ , 1.0 g  $\text{H}_3\text{BO}_3$ , 1.0 g KI and 15 g  $\text{Na}_2\text{EDTA}$ .

Solution of vitamins (500 ml): 25 mg d-biotin is dissolved in 10 mL 0.1 M NaOH + approx. 400 mL  $\text{H}_2\text{O}$ . Adjust the pH to 6.5 with HCl. 500 mg Ca-Pantothenat, 500 mg Thiamin-HCl, 500 mg Pyridoxin-HCl, 500 mg Nicotinic acid and 100 mg p-aminobenzoic acid. 12.5 g m-Inositol was added and pH was adjusted to 6.5 and the solution was filter sterilize. (Modified from the medium described in Verduyn *et al.*, 1992).

Cultivation was performed in baffled shake flasks at 30 °C with vigorous shaking.

### 5.2.3 Experimental setup

A single colony was inoculated in a pre-warmed CBS medium. The pre-pre-culture was inoculated in 6-8 hours 30 °C. A dilution series was made from this pre-pre-culture and inoculated overnight. Pre-warmed CBS medium was inoculated with a pre-culture, from the dilution series, with an  $\text{OD}_{600}$  of 0.2 – resulting in an exponentially growing culture – a balanced growing culture as the exponential growth have lasted for more than eight generations. The balanced culture was compared to a stationary phase overnight culture propagated in CBS medium.

### 5.2.4 Fluorescence microscopy, spectrofluorophotometry and flow cytometry

The spectrofluorophotometer (RF-5301PC Shimadzu) was used to measure fluorescence intensity on a population level. The emission and excitation slit were set to 5. The excitation wavelength was set at 488 nm and the emission spectrum was measured. Emission intensity was determined at 510 nm.

The control strains FE522 exhibited a measurable, but still close to the detection limited of the spectrofluorophotometer, autofluorescence, which increased during the cultivation due to increase in cell density. Due to the fact that the autofluorescence of the cells was close to (and often below) the detection limited of the apparatus, it was not possible to include in the determination of the GFP expression.

A FACSCalibur (BD Biosciences), with a 15 mW, 488 nm, air-cooled argon-ion laser was used for single cell analysis. The flow cytometry settings used for analysis of *S. cerevisiae* were FSC E-1, SSC 319 V, FL1 623 V (all logarithmically amplified). A side scatter threshold was applied to gate out much of the noise (at channel 72), a threshold on forward scatter at channel 108 was also included.

During acquisition of *S. cerevisiae* all events were saved until the software counted 50,000 events of interest (gated fluorescent events). The samples were resuspended in PBS solution (0.137M NaCl, 0.003M KCl, 0.012M NaHPO<sub>4</sub> and 0.002M KH<sub>2</sub>PO<sub>4</sub> pH 7.2) and placed on ice before flow cytometry analysis.

Histogram gates were applied in the post-processing of the flow cytometer data according to fluorescence. Histogram gating was done using Flowing Software (v2.5.1, Turku Centre for Biotechnology, Turku, Finland), before statistical analysis was done in JMP®. FlowJo® was used for graphical display of the data (contour plots, overlay and offset histograms).

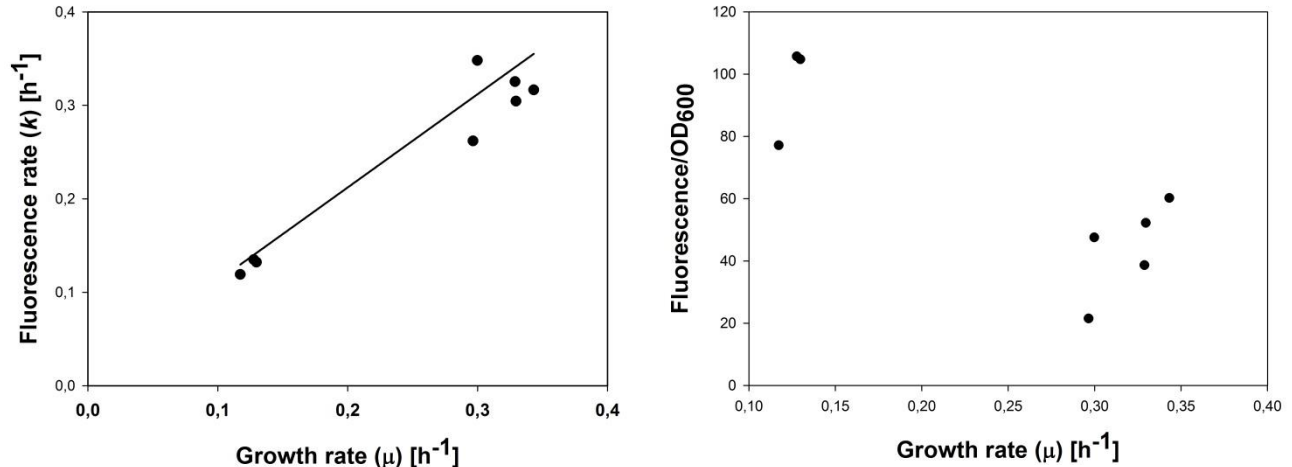
## 5.3 Results and discussion

### 5.3.1 *S. cerevisiae* GFP reporter strain expressed growth rate regulated fluorescence on a population level

FE440 was initially validated on a population level to determine if GFP expression was growth rate regulated. Spectrofluorometric analysis for determination of GFP fluorescence on a population level has not been published for FE440. FE440 and the control strain FE522 were cultivated in minimal CBS medium with different carbon sources (glucose, galactose and maltose). Cell density and fluorescence was determined by spectrometry and spectrofluorometry, respectively.

The expression of GFP in the reporter strain did not affect the growth rate, when compared to the non-fluorescent control strain (FE522) (Carlquist *et al.*, 2012) – which is in accordance with what I observed (data not shown).

The linear correlation between specific fluorescence rate ( $k$ , h<sup>-1</sup>) and growth rate is shown in the scatter plot in Figure 20(left). The specific fluorescence (fluorescence per OD) was plotted against the growth rate, see Figure 20(right). A linear relation could not be observed, indicating that the GFP expression in the yeast cells was not growth rate regulated. Thus the GFP expression in strain 440 approximately correlated with the growth rate, but it was not growth rate regulated.



**Figure 20 (left)** Specific fluorescence rate  $k$  was positive correlated to growth rate. **(right)** Specific fluorescence (fluorescence per OD) did not correlate to growth rate. To generate the data used in the scatter plots were the fluorescence and optical density measured for the green fluorescent *S. cerevisiae* reporter strain FE440 when propagated in chemically defined medium supplemented with galactose, maltose or glucose.

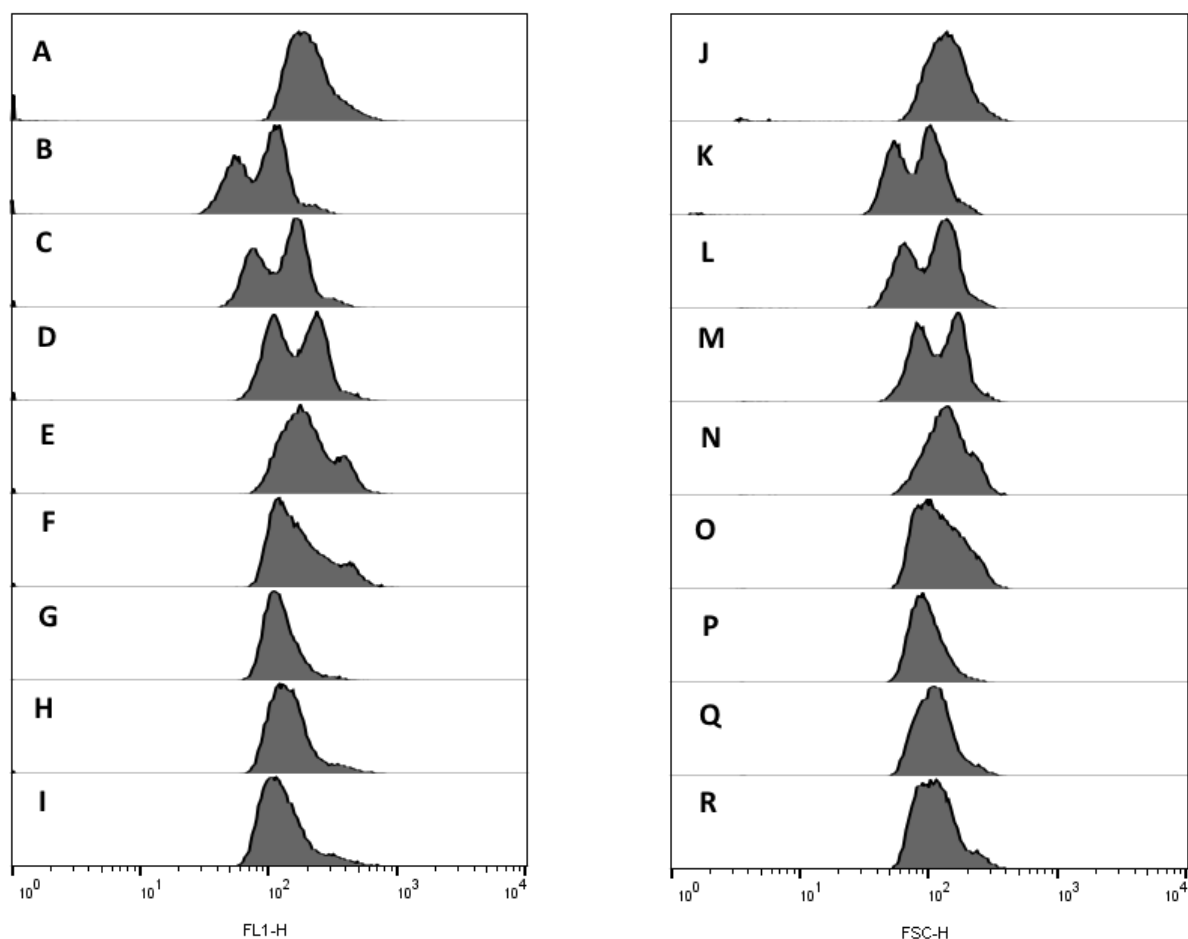
### 5.3.2 Large standard deviation in single cell growth rates for the exponentially growing culture

With the aim to study population heterogeneity in *S. cerevisiae* cultures inoculated with either a stationary phase pre-culture or an exponentially growing pre-culture samples from shake flask experiments were analyzed on a single cell level by flow cytometry. The experiment was performed in duplicate; the results presented below are from replica A, results for replica B can be found in section 11.4.1.2. When analyzing the results the influence of an introduced source of error needs to be included: The stationary phase pre-cultures were diluted in a ratio 1:20, resulting in an OD<sub>600</sub> of  $\sim 0.5$ . The exponential growing pre-cultures, having an OD<sub>600</sub> of 0.15 and 0.35, respectively were diluted in a ratio of 1:10, resulting in a considerable lower cell concentration.

Due to the more complex size distribution, compared to bacteria, information of cell size in the culture can be important to include in the analysis of single cell growth rate. In *S. cerevisiae* is the cell cycle progression coupled to cell growth. The initiation of the yeast cell cycle at “start” (the regulatory step in the mitotic cell cycle) occurs during the transition period between G1- and S-phase. The yeast cell has attained a critical size at “start” and a small bud emerges (Hartwell & Unger, 1977; Jagadish & Carter, 1977; Jagadish *et al.*, 1977). A mother cell remains relative unaltered during growth of the bud (Johnson, 1965). The cell size, which depends on the growth rate, affects the duration of G1, smaller cells are delayed in the entry of the S-phase (DNA replication) (Jagadish *et al.*, 1977; Tyson *et al.*, 1979).

Cells arrested in stationary phase have exited the cell cycle thus entering G<sub>0</sub>. Cell in G<sub>0</sub> has not reached the regulatory step (start) and is in a unbudded, prereplicative status (Reed, 1980).

When analyzing the flow cytometry data a multimodal distribution during the lag phase of the cultivation inoculated with the stationary phase culture was observed. The multimodal distribution (can be visualized using a 3-D terrain plot, data not shown) is shown below in the off-set histograms in Figure 21



**Figure 21** Off-set histograms showing the distribution of fluorescence per cell [FL1-H] (left) and forward scatter per cell [FSC-H] (right). (A and J) the distributions observed for the balanced growing culture taken three generations after inoculation, growth curve can be seen above the box-plot in Figure Figure 22B. (B and K) the sample taken 10 minutes after inoculation with a stationary phase pre-culture – the growth curve for this culture can be seen above the box-plots in Figure Figure 22A. (C and L) distributions observed in the sample taken 60 minutes after inoculation. (D and M) 126 minutes after inoculation. (E and N) 191 minutes after inoculation. (F and O) 223 minutes after inoculation. (G and P) 258 minutes after inoculation. (H and Q) 291 minutes after inoculation. (I and R) 325 minutes after inoculation.

It can be seen from the frequency distributions of the fluorescence per cell [FL1-H] presented as off-set histograms in Figure 21 A-I, that several subpopulations were present during this phase. Comparing with the distribution of cell size (forward scatter per cell Figure 21J-R) similar subpopulations can be seen. The subpopulation observed in the forward scatter per cell during lag phase could possibly be cause by cells re-

entering the cell cycle after having being arrested in  $G_0$ . As the cells was growing slower in the culture inoculated with a stationary phase pre-culture the cells could potentially be a little smaller than the cells in the balanced culture. – this can also be observed from the forward scatter measurements (see section 11.4)

The sample variation, considering both fluorescence per cell and forward scatter per cell, observed in the balanced culture was imperceptible. The sample taken three generations after inoculation was thus used as representative for the balanced culture (Figure 21A and J).

Using the distribution of the specific cellular fluorescence as an indication of distribution of single cell growth rates only little variation between the samples can be observed for the balanced culture. For the culture inoculated with a stationary phase pre-culture a larger variation between the box-plots was observed (Figure 22A). The CV as well as mean to median ratio indicated that the culture inoculated with the stationary phase culture was less heterogeneous than the culture inoculated with an exponential growing pre-culture.

The amount of stable RNA in yeast cells depends on their growth rate: fast-growing cells have a much higher RNA content than slow-growing cells (Boehlke & Friesen, 1975; Wehr & Parks, 1969). It was thus expected that the cellular fluorescence of stationary phase cells once inoculated in fresh medium would increase during the lag phase, as the activity of the ribosomal promoter would increase. The increase rate of GFP expression combined with that the cell was not dividing yet increased the total cellular fluorescence of the cell (see Figure 61).

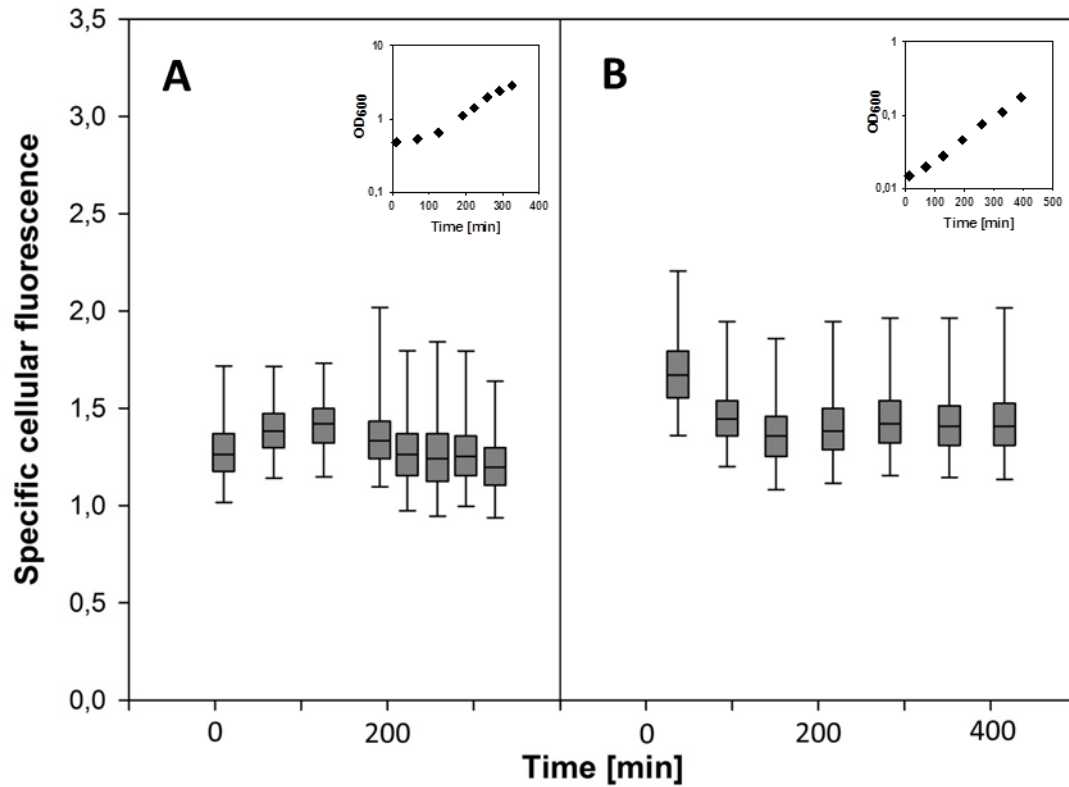
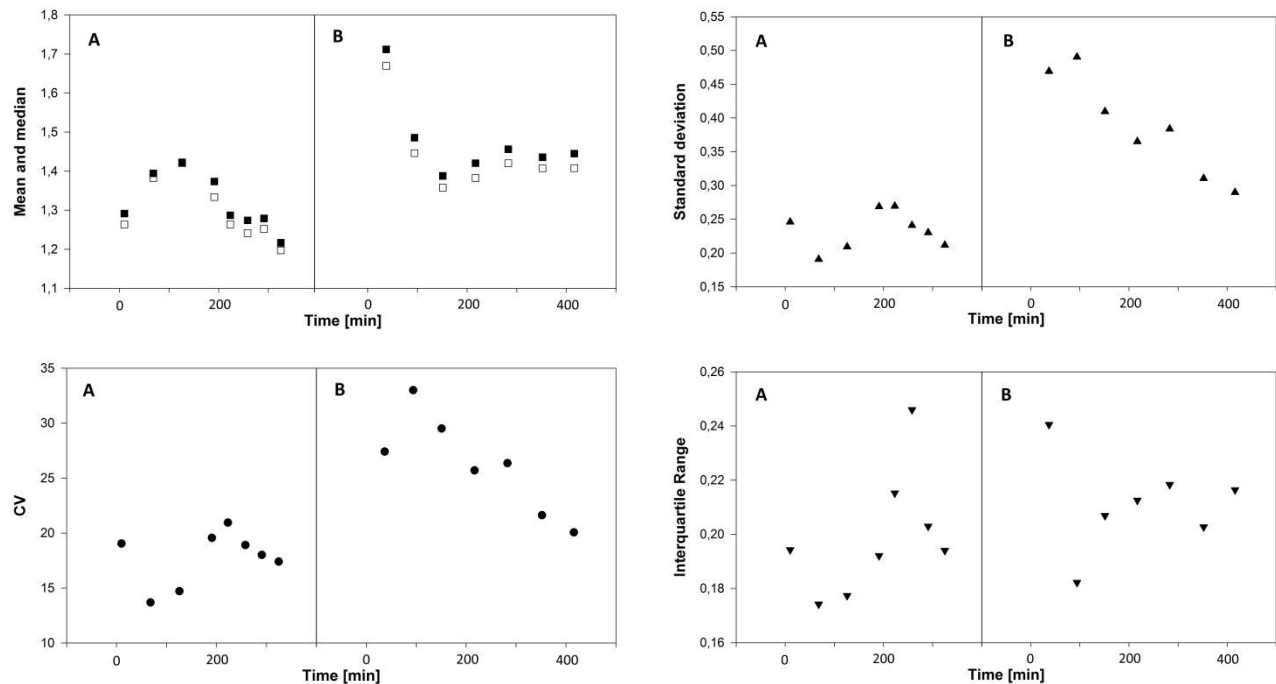


Figure 22 Box-plot showing the distribution of specific cellular fluorescence (single cell growth rate) during cultivation with the yeast reporter strain FE440 cultivated in chemically defined medium inoculated with either a stationary phase overnight culture (A) or an exponentially growing culture (B). The growth curves of the culture are shown above the box-plots, each diamond symbol in the growth curves can be linked to a box-plot. Growth rate in exponential phase in CBS medium: (A)  $0.29 \text{ h}^{-1}$  ( $R^2=0.9789$ ) and (B)  $0.35 \text{ h}^{-1}$  ( $R^2=0.9982$ ).



**Figure 23** Descriptive statistics based on specific cellular fluorescence measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. The associated growth curves are shown in Figure 22.

The descriptive statistics shown in Figure 23 show that the standard deviation and the coefficient of variation were lower for the culture inoculated with a stationary phase culture. Could the lower level of heterogeneity in the culture inoculated with the stationary phase overnight culture be caused by that the overnight culture had not reached stationary phase before being used? If the culture had not entered stationary phase the culture would upon inoculation in fresh medium be in a quasi-balanced state, as the culture definitely must have been in the late exponential phase after 16 hours of cultivation. A paper published in 2015 presents a growth curve based on fermentation data for a *S. cerevisiae* GFP expressing strain cultivated in shake flask with CBS 2 % glucose medium. The culture enters stationary phase somewhere between 12 hours and 24 hours after inoculation. The diauxic shift, detected by expression of GFP regulated by a promoter strongly repressed by glucose, occurred after 12 hours as the glucose was consumed (determined by HPLC measurements) (Williams *et al.*, 2015). The stationary phase pre-culture used in this study was made by inoculating a single colony from an YNB plate in 10 ml chemically defined medium (CBS, 2 % glucose). The cell in the 10 mL medium were propagated overnight (>16 hours), hence the pre-culture must had passed the diauxic shift and most likely had reached stationary phase before use.



yEGFP (yeast enhanced GFP based on the GFPmut3 chromophore mutation) was used as reporter in FE440 (Carlquist *et al.*, 2012; Cormack *et al.*, 1996, 1997). It is known that acidification of the medium can affect the fluorescence intensity of GFP (section 4.1 and Campbell & Choy, 2001) and by lowering the medium pH a source of error might be introduced. It has been shown that *S. cerevisiae* acidifies the medium, when propagated in minimal (Verduyn) medium the pH of an overnight culture was measured to be 4.1 (Delvigne *et al.*, 2014). To avoid decrease in pH a buffer (MES) was added to the chemically defined medium to prevent acidification and during propagation of the cells, the pH was kept above 6 (data not shown).

So assuming that the pre-cultures (a stationary inoculum was also made for replica B) had reached stationary phase before use and knowing that the pH of the medium at all time was kept above 6 the experimental setup should work as intended, conclusively the lowest level of heterogeneity can be observed in the cultures inoculated with a stationary phase pre-cultures (the results for replica B can be seen in section 11.4.1.2, it shows the same tendency as replica A but not as pronounced).

Considering fluorescence per cell and forward scatter per cell it was observed that the variation between the sampled was lowest for the balanced growing culture. The doubling time for the cultures presented in replica A was, during exponential phase, 145 minutes (stationary phase inoculum) and 119 minutes (exponential growing inoculum) for the cultures, respectively. The slower growth rate for the culture inoculated with the stationary phase pre-culture indicates that the culture was not growing at a maximum growth rate, and due to the high start OD it may be assumed that the culture prepared for stationary phase.

.

## 6 Thermal stress introduces phenotypic variations in *B. subtilis*

### 6.1 Introduction

Results based on analysis of the *B. subtilis* reporter strain AE099, expressing GFP under regulation of a ribosomal promoter, showed that population heterogeneity was affected by the choice of medium and the history of the inoculum (see submitted manuscript included in section 3).

Wanting to investigate if population heterogeneity would increase if a culture of genetically identical cells were exposed to stress, flow cytometry single cell analysis was made with a culture of the *B. subtilis* reporter strain AE099. The samples for analysis were taken from a growth experiment where the cells were propagated as a balanced growing culture at 37°C and then moved to a higher incubation temperature, exposing the cells to an abrupt temperature increase.

Thermal stress was expected to have an effect on the population heterogeneity as heat shock (an abrupt increase in temperature) activates transcription of heat shock proteins in bacteria. Two main classes of protein is being synthesized; chaperones and adenosine tri-phosphate (ATP)–dependent proteases. Chaperones ensure that polypeptides will fold or assemble properly in the cell (Georgopoulos & Welch, 1993). Proteases will degrade the misfolded proteins, which are unable to refold into their native 3-dimensional structure (Gottesman, 1996). When a cell is exposed to thermal stress the major problem is the immediate appearance of denaturated and misfolded proteins, which tend to aggregate. The chaperones and ATP-dependent proteases will recognize and degrade these proteins (Gottesman *et al.*, 1997).

Once exposed to heat shock or thermal stress will the synthesis of heat shock genes, which in *B. subtilis* counts more than 200, be activated. The activation is primarily induced by increase in transcriptional activity (Schumann, 2003). Heat shock proteins are under regulation of different sigma factors. In *B. subtilis* are the heat shock proteins which are regulated by the secondary sigma factor (sigma-B) the general stress proteins. The sigma-B factor regulated stress genes are also expressed during starvation, salt stress, and when the cells is exposed to ethanol (Hecker *et al.*, 1996).

In the cells two regulatory principles exists: direct and indirect heat sensing. RNA or protein molecules that change conformation depending on the temperature are direct heat sensors. The indirect heat sensors are chaperons. The heat shock genes controlled by direct sensors are expressed at a high rate as long as the cells are exposed to the increased temperature, the genes controlled by the indirect sensors are induced momentarily (Schumann, 2003).

Thermal stress was expected to result in extensive changes in the gene expression pattern and thus increase in both intrinsic and extrinsic noise leading to increased population heterogeneity. When using a GFP reporter strain to study how thermal stress affects bacteria cells, it should be taken into interpretation that the increased temperature might affect the fluorescence of the cells, but how much has not been determined and will probably depend on the GFP variant and the analysis method; wtGFP is documented to be quite thermostable (Bokman & Ward, 1981), Vessoni Penna et al determines that the concentration of GFPuv

decrease 11 % at 45 °C, relative to initial native protein concentration determined from a standard curve (Vessoni Penna *et al.*, 2004). Simering *et al.* observed that the fluorescence intensity of the wtGFP was 7.4 fold lower when cultivated at 42 °C compared to at 37°C (Siemering *et al.*, 1996).

## **6.2 Materials and Methods**

### **6.2.1 Media**

*B. subtilis* was grown in Spizizen minimal salt medium supplemented with 1 mg of thiamine and 40 mg tryptophan per liter, as previously described (Saxild & Nygaard, 1987), this medium will be referred to as chemically defined medium in this section. Glucose was added to a final concentration of 0.4 %. This medium will in this section be referred to as chemically defined medium. The strains were cultivated in Erlenmeyer flasks at 37°C with vigorous shaking.

### **6.2.2 Fluorescence microscopy, spectrofluorophotometry and flow cytometry**

Fluorescence measurements at culture levels were carried out with a Shimadzu RF-5301PC fluorometer (Columbia, MD). GFP excitation was performed at 488 nm, and emission was monitored at 510 nm with excitation and emission slit widths of 5:5. The fluorescence of *B. subtilis* reporter strain and control strain was determined directly in the minimal medium.

Fluorescence microscopy was done using a Zeiss Axioplan fluorescence microscope, GFP fluorescence was analyzed with the use of Semrock GFP-3035C BrightLine single-band filter set. Fluorescence microscopy was used to isolate transformants.

For flow cytometric analysis were the cells diluted in PBS solution: NaCl, 0.137 M; KCl, 3 mM; NaHPO<sub>4</sub>, 12 mM; and KH<sub>2</sub>PO<sub>4</sub>, 2 mM; pH 7.4, and measured directly on a BD FACSCalibur (Becton-Dickinson, NJ, USA) flow cytometer. Fluorescence emission levels were measured using a band pass filter at 530/30 nm (FITC). For each sample approximately 100,000 cells were analyzed.

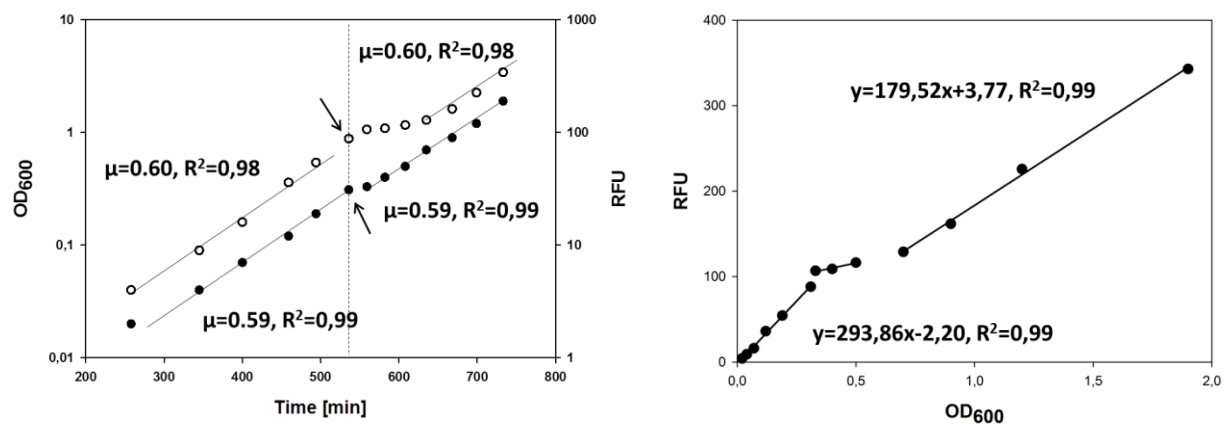
## **6.3 Results and Discussion**

### **6.3.1 Growth proceeded at similar rate after exposure to continuously thermal stress**

Before the green fluorescent *B. subtilis* reporter strain AE099 was analyzed on a single cell level to investigate population heterogeneity in a culture exposed to thermal stress, a population level analysis was performed. The population analysis would establish how growth rate and the growth regulated fluorescence would respond to the temperature increase. A culture propagated in chemically defined medium was cultivated as a balanced growing culture at 37°C, when the optical density was 0.3 the balanced growing culture was moving to a 48°C hot water bath. It was observed that the biomass stopped increasing once the

culture was exposed to thermal stress. The biomass increase stagnated for approximately an half an hour. After the stagnation the biomass increased exponential, at a similar rate as before the temperature increase. The fluorescence was, when the culture was growing balanced, increasing at the same rate as the biomass. The exposure to thermal stress made the fluorescence intensity stagnates for approximately 100 minutes for subsequently to increase again, at a rate similar as the one before the temperature shift (Figure 24, left). The fluorescence per biomass decreased after the culture had adapted to 48°C (Figure 24, right). This could indicate that the cells grew slower or that the fluorescent proteins were affected by the increased temperature.

The experimental setup was copied and this time the cells were analyzed by on a single cell level by flow cytometry.

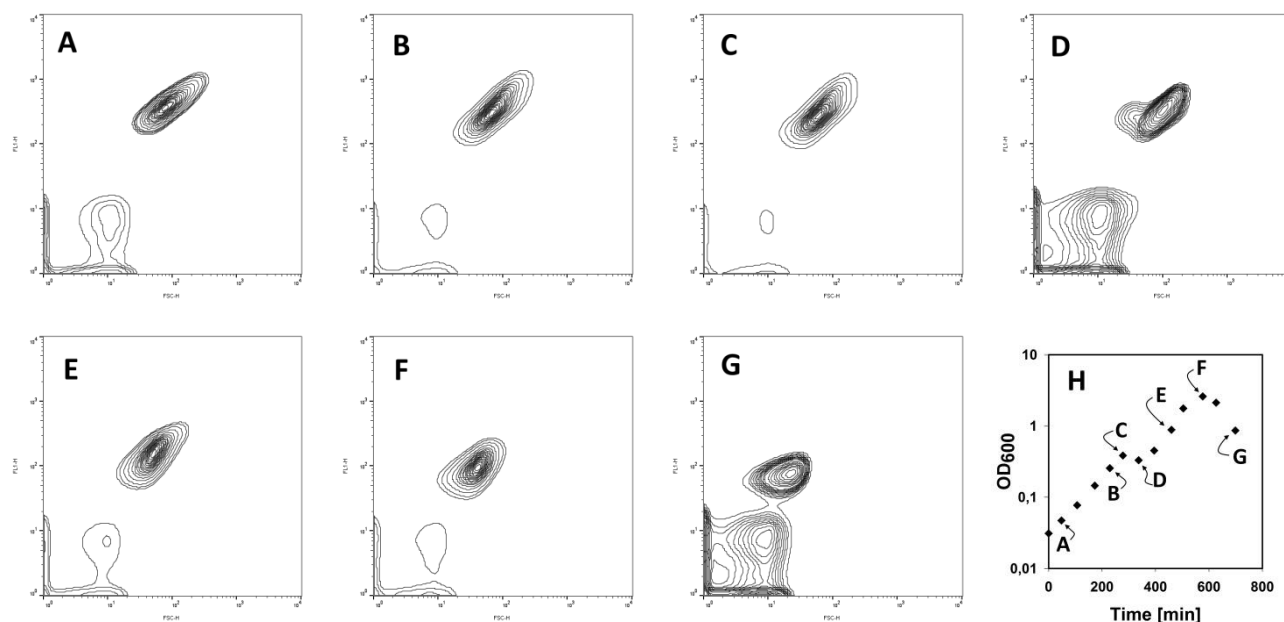


**Figure 24 (left)** Growth rate and fluorescence rate measured for a growing *B. subtilis* culture before and after exposure to thermal stress. Fluorescence rate is the increase in relative fluorescence per time unit. The reporter strain AE099 was propagated at 37°C. The culture was exposed to thermal stress (48°C) 536 min after inoculation. The measurements taken immediately after the culture was incubated at 48°C are marked by arrows and a dashed line. Growth rate and fluorescence rate plateaued immediately after the temperature shift. The growth rate and fluorescence rate increased after the stalling at rates similar to the once observed before the temperature shift. The growth rates and fluorescence rates are shown in the the plot next to the lines showing which measurement were included in the determination of the rate. Closed symbols: Optical density (OD<sub>600</sub>). Open symbols: Relative fluorescence units (RFU).

**(right)** Fluorescence per biomass measured before and after the abrupt temperature increase. The equation for the regression lines are shown in the plot. A decrease in fluorescence per biomass was observed after introduction of thermal stress.

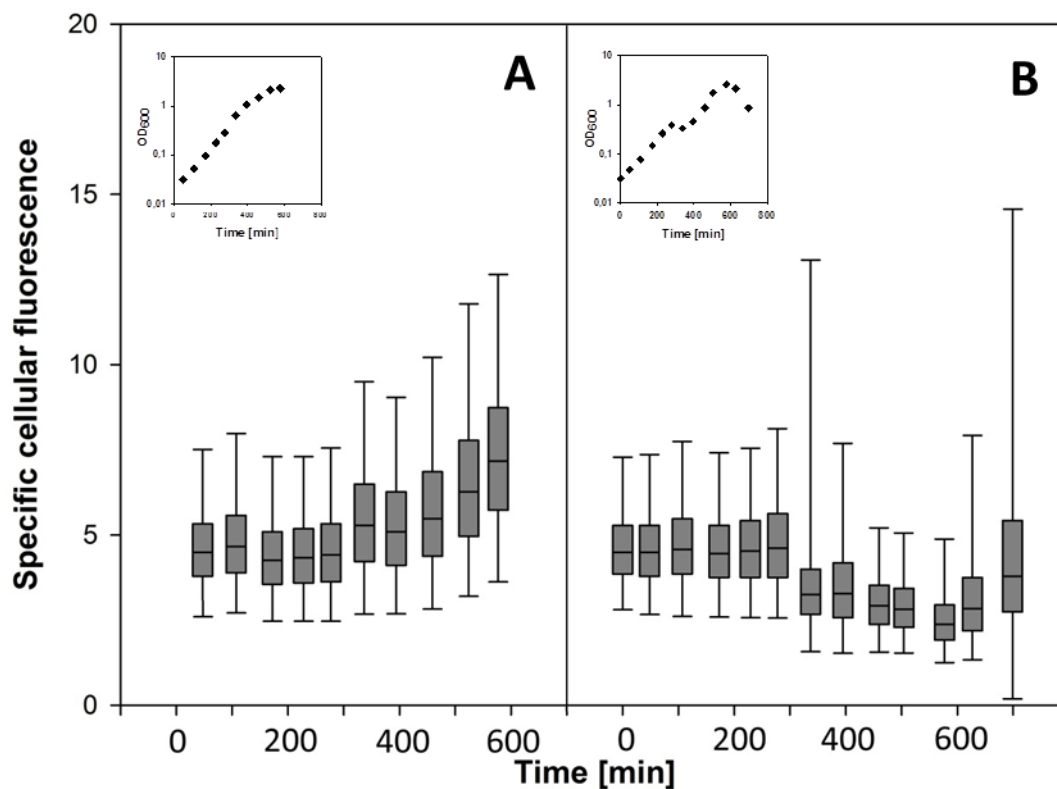
### 6.3.2 Thermal stress increases population heterogeneity

Flow cytometry analysis of a *B. subtilis* culture exposed to a temperature shift from 37°C to 48°C was made to investigate how thermal stress affected population heterogeneity. The results of the flow cytometry analysis are shown in contour plots in Figure 25.



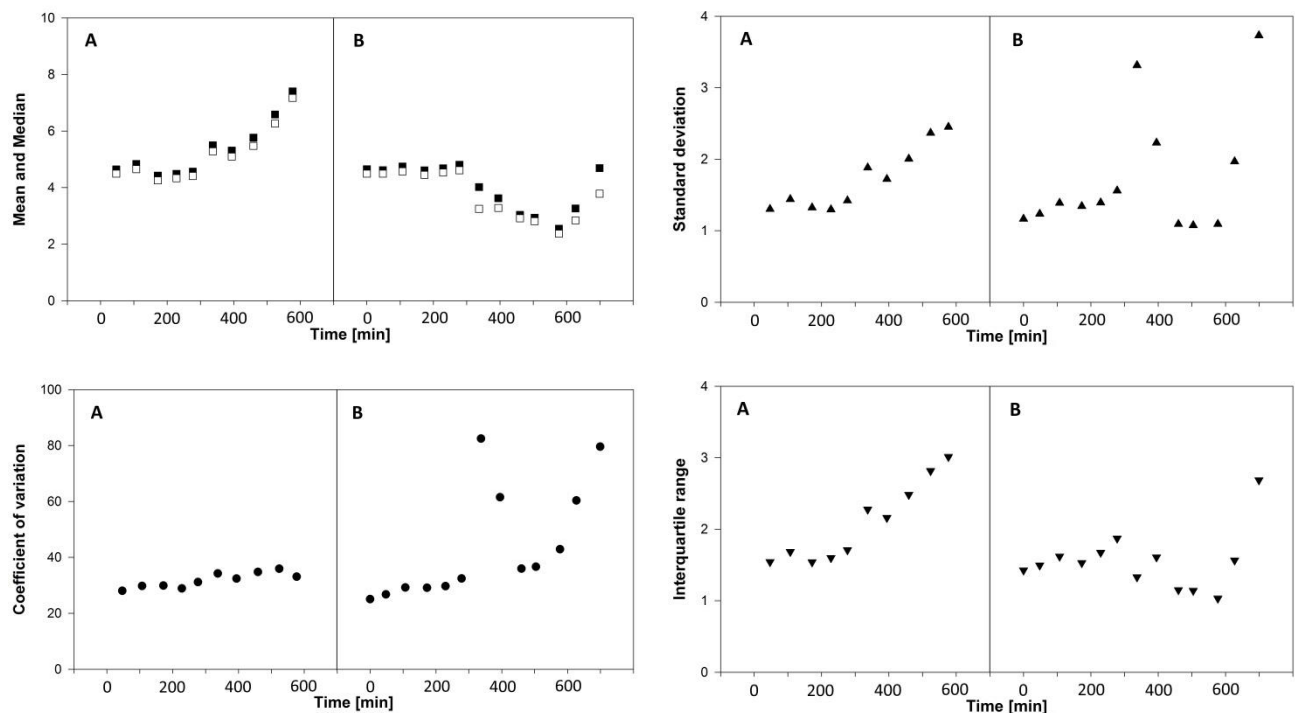
**Figure 25** Contour plot showing forward scatter per cell (horizontal axis) and fluorescence per cell (vertical axis) for the *B. subtilis* reporter strain AE099 cultivated as a balanced culture at 37°C (A and B) and after the inoculation temperature was increase to 48°C (C-G). The growth curve can be seen as scatter plot in (H). In the scatter plot is marked when the samples shown in contour plot A – G were taken. 270 minutes after inoculation the culture was exposed to an abrupt increase in incubation temperature (from 37°C to 48°C), the sample taken few minutes after the shift is shown in (C).

The population profiles gradually changed after the incubation temperature was increased. The cells gradually became smaller and less fluorescent once exposed to thermal stress. The fluorescent cells were gated and the specific cellular fluorescence was computed as described in section 2. Box-plot showing the distribution of the specific cellular fluorescence for a balanced growing culture at 37°C can be seen in Figure 26A, box-plot showing the distribution of the culture propagated as a balanced growing culture at 37°C and then moved to 48°C can be seen in Figure 26B. (Box-plot showing the variation in fluorescence per cell and forward scatter per cell can be found in supplementary materials section 11.5).



**Figure 26** Box-plot showing the variation in the specific cellular fluorescence measured for two cultures of *B. subtilis* strain AE099. The balanced growing culture propagated at 37°C was included as a control (A). A balanced growing culture propagated at 37°C was 278 minutes after inoculation exposed to thermal stress by increasing the incubation temperature to 48°C (B). The diamond symbols in the growth curves shown above the box-plots indicate when the samples for flow cytometry analysis were harvested.

The distribution of single cell growth rate was for the balanced growing culture (Figure 26A) constant until late exponential phase where the heterogeneity started to increase gradually. The distribution of single cell fluorescence growth rate for the culture exposed to temperature shift is shown as box-plot in Figure 26B. Here a constant level of heterogeneity within the samples can be observed when the culture was growing balanced at 37°C. The sixth box-plot shows the distribution of specific cellular fluorescence immediately after the temperature shift. In the seventh box-plot in Figure 26B is the distribution of specific cellular fluorescence little more than an hour after the abrupt temperature increase shown. As can be observed for the descriptive statistics shown in Figure 25 did the specific cellular fluorescence mean decrease at this time (an hour after the temperature increase), but the standard deviation and the coefficient of variation increased dramatically, indicating that the population heterogeneity increased upon thermal stress.



**Figure 27** Descriptive statistics for the specific cellular fluorescence measured for the *B. subtilis* reporter strain. (A) Balanced growing culture propagated at 37°C. (B) Balanced growing culture propagated at 37°C and 278 minutes after inoculation exposed to thermal stress to 48°C (Top left) mean: black squares and median: white squares. (Top right) standard deviation: triangles. (Bottom left) coefficient of variation: diamonds. (Bottom right) interquartile range: triangle, down. The descriptive statistics can be linked to the growth curves presented in Figure 26.

The mean and median values determined for the specific cellular fluorescence decreased gradually as the culture was exposed to thermal stress. The decrease in specific cellular fluorescence could indicate that the cells grew slower after being incubated at 48°C. Looking at the distribution of forward scatter per cell and fluorescence per cell it can be observed that the interquartile range and the cell size mean value increased after the temperature shift, whereas the fluorescence level at first was unaffected but after 100 minutes at the increased temperature it decrease dramatically.

Combining the observations done on a single cell level with the observations done on a population level interesting question emerges. On a population level it was determined that the biomass increased and then due to the heat shock stagnated shortly for subsequently to increase at the same rate ( $\mu_{37^{\circ}\text{C}} 0.57 \text{ h}^{-1}$  and  $\mu_{48^{\circ}\text{C}} 0.61 \text{ h}^{-1}$ ). However, the cells size increased in the stagnation period for subsequently to decrease, the mean values can be seen in Figure 70, top left, A. This could indicate that once the incubation temperature raised some cells died while others increase in biomass, and when they larger cell then divided the cell became smaller, but they would have an increased growth rate.

It would be difficult to conclude whether single cell growth rate decreased after the temperature shift or whether the decreased was caused by the fact that the fluorescence of GFP decreases at elevated temperature, as the chromophore has difficulties to mature at warmer temperatures (Tsien, 1998). GFP that have matured at lower temperature will stay fluorescent once moved to temperatures up to 65 °C (Bokman & Ward, 1981), thus the GFP which has matures before the reporter strain was exposed to thermal stress would most likely still be fluorescent once moved to 48°C. Some of the immature GFP become fluorescent would probably also mature at the elevated temperature, but how large a percentage is unknown.

The increase in incubation temperature affected the population profiles (see contour plots in Figure 25). On a population level were the growth rate and fluorescence rate determined and a stalling after the heat shock was observed for both parameters. The population heterogeneity determined by standard deviation and coefficient of variation for specific cellular fluorescence increased after the shift, for again (after more than 100 minutes) to drop to the level observed during balanced growth.

From the results it could be observed that thermal stress affected the population profile and transient increased population heterogeneity.



## 7 Single analysis of a *E. coli* strain expressing GFP under regulation of an inducible promoter

### 7.1 Introduction

Fluorescent proteins has in several publication been used to analyze cell-to-cell variability introduced by promoter activity (intrinsic noise) (Bar-Even *et al.*, 2006; Goulian & van der Woude, 2006; Piersma *et al.*, 2013). Intrinsic noise is defined as the part of the total cellular noise arising from gene expression (Elowitz *et al.*, 2002). Goulian and van der Woude showed in a paper from 2006 results from an analysis of an *Agrobacterium tumefaciens* strain with GFP integrated in *virE*. The promoter regulating the transcription of *virE* is an acetosyringone (AS) inducible promoter. Cultures of the green fluorescent *A. tumefaciens* were propagated with different concentrations of AS, and frequency distribution (determined by microscopy) clearly indicates cell-to-cell variability in transcription of the VirA/VirG-regulated gene for intermediate concentrations of AS (Goulian & van der Woude, 2006).

An *E. coli* strain with GFP integrated in *lacZ* (MG1655 *lacZ::gfp*) was also constructed by Goulian and van der Woude. This strain was used for analysis of fluorescence intensity at a population level at different concentrations of the substrate analogue Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and the analysis showed that the promoter was induced at different levels dependent on the IPTG concentration (Goulian & van der Woude, 2006). The *E. coli* MG1655 *lacZ::gfp* reporter strain was, upon enquiry, received <sup>1</sup>.

The MG1655 used for the construct (MG1655 *lacZ::gfp*) is *lacY*, which is considered important when observing *lacZ* promoter activity, as the induction of the *lacZ* promoter is a so-called all-or-none enzyme induction (Novick & Weiner, 1957). If IPTG is present at higher concentrations the lac operon is fully de-repressed and cells express the transmembrane symporter (lactose permease, encoded by *lacY*) and the promoter remain highly activated. At low concentrations, however, cells that were previously unindicted and do not have any permease in their membranes do not respond to the low level of IPTG and the *lacZ* promoter is not induced. Cells that were previously induced and still have some permease remains activated by lower level of IPTG and hence the promoter transcribes the lac operon (Veening *et al.*, 2008). If lactose permease is not expressed (the strain is *lacY*) will activation of the lac operon depend only on passive diffusion of IPTG and hence the promoter will be more gradually activated than in a *lacY*<sup>+</sup> strain (Fernández-Castané *et al.*, 2012; Jensen *et al.*, 1993).

The main objective by requesting the *E. coli* MG1655 *lacZ::gfp* strain was to investigate cell-to-cell variability on a single cell level, but initially the GFP expression at various IPTG concentrations was analyzed on a population level, in an attempt to replicate the result published in the paper by Goulian and van der Woude.

---

<sup>1</sup> Thanks to Marjan van der Woude for providing the *E. coli* MG1655 *lacZ::gfp* strain (Goulian & van der Woude, 2006).

## 7.2 Materials and methods – *E. coli*

### 7.2.1 Strains

*E. coli* MG1655 *lacZ::gfpmut3.1 lacY* was a kind gift from Marjan von der Woude (Goulian & van der Woude, 2006). This strain will be referred to as MG1655 *lacZ::gfp*.

*E. coli* strain 2670 (MG1655) was used as control strain.

### 7.2.2 Medium

*E. coli* was for analysis of GFP expression under regulation of the *LacZ* promoter propagated in AB medium (Clark & Maaløe, 1967) supplemented with 0.2 % glycerol and 10  $\mu$ g/ml uracil. This medium will in this section be referred to as chemically defined medium.

Cultivation was performed in baffled shake flasks at 37°C with vigorous shaking.

### 7.2.3 Spectrofluorophotometry and flow cytometry

Fluorescence intensity on a population level was measured with a RF-5301PC Shimadzu spectrofluorophotometer. The emission and excitation slit were set to 5. The excitation wavelength was set at 488 nm and the emission spectrum was measured. Emission was determined at 510 nm. The fluorescence of the cells was determined in the growth medium.

Chloramphenicol (CAM) was immediately added to the samples taken for spectrofluorimetric analysis (final concentration of CAM 100  $\mu$ g/ml). The CAM treated samples were placed on ice before analysis.

A FACSCalibur (BD Biosciences), with a 15 mW, 488 nm, air-cooled argon-ion laser was used for single cell analysis. The flow cytometry settings were: FSC E02, SSC 381 V, FL1 831 V (all logarithmically amplified). A side scatter threshold was applied to gate out much of the noise (at channel 199).

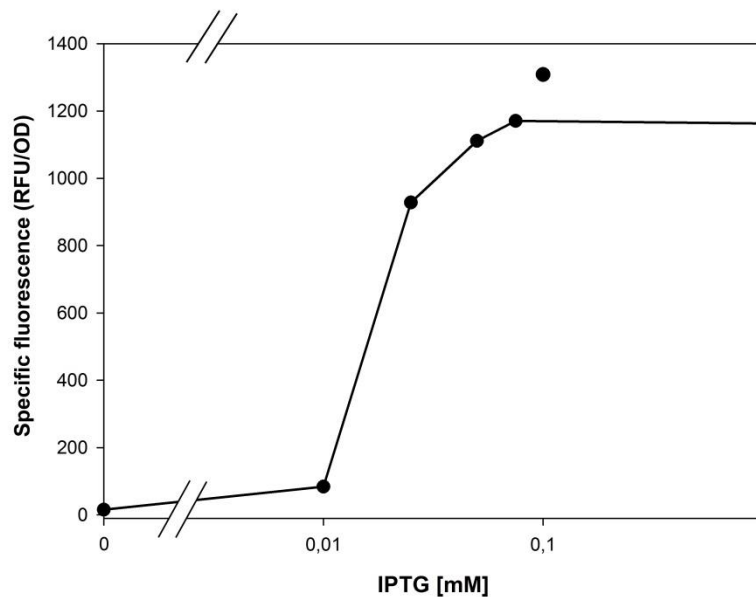
During acquisition of *E. coli*, all events were saved until the software counted 100,000 events. The samples were centrifuged, resuspended in PBS solution (0.137M NaCl, 0.003M KCl, 0.012M NaHPO<sub>4</sub> and 0.002M KH<sub>2</sub>PO<sub>4</sub> pH 7.2) and placed on ice before being analyzed by flow cytometry.

FlowJo was used for graphical display of the data.

## 7.3 Results and discussion

### 7.3.1 IPTG induction of the *lacZ* promoter observed on a population level

The reporter strain MG1655 *lacZ::gfp* was studied on a population level to observe GFP expression at different IPTG concentrations. An experiment was conducted to replicate the result presented in the paper by Goulian and van der Woude; IPTG was added to balanced growing cultures of MG1655 *lacZ::gfp* and the GFP expression was determined several times during the experiment (Goulian & van der Woude, 2006). In the *E. coli* reporter strain was the GFP expression, at intermediate IPTG concentrations, lower compare to when the promoter was fully induced (Figure 28). This observation was in agreement with the result published by Goulian and van der Woude.



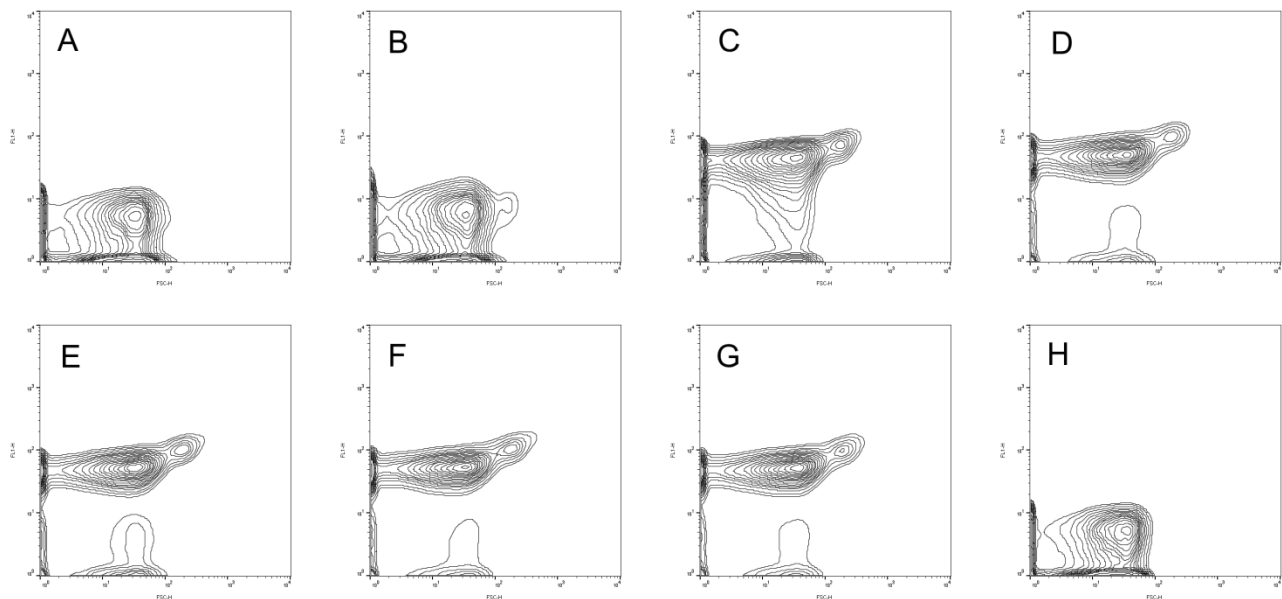
**Figure 28** GFP expression (fluorescence per OD) measured at different IPTG concentrations for the reporter strain MG1655 *lacZ::gfp* cultivated in chemically defined medium. The autofluorescence from the cell (determined from measurements of MG1655 *lacZ::gfp* cultivated in medium without IPTG, and measurements of 2670 cultivated in the presence of 1 mM IPTG) was subtracted before the positive correlation between RFU and OD was determined. The slope of the linear dependency between fluorescence and optical density is in the graph above plotted at various IPTG concentrations. Specific fluorescence measured for the cultivations made with 0.1 mM IPTG in the medium was not included in the tendency line.

The result presented in the paper by Goulian and Woude was reproducible; GFP expression in the MG1655 *lacZ::gfp* strain was increased at higher IPTG concentrations, reaching a maximum expression at 0.1 mM IPTG. The doubling time for MG1655 *lacZ::gfp* in AB medium supplemented with glycerol was 132 minutes  $\pm$  5.4 minutes – the cells had been growing exponentially for more than two generations after induction with IPTG, before the first spectrofluorimetric measurements were made.

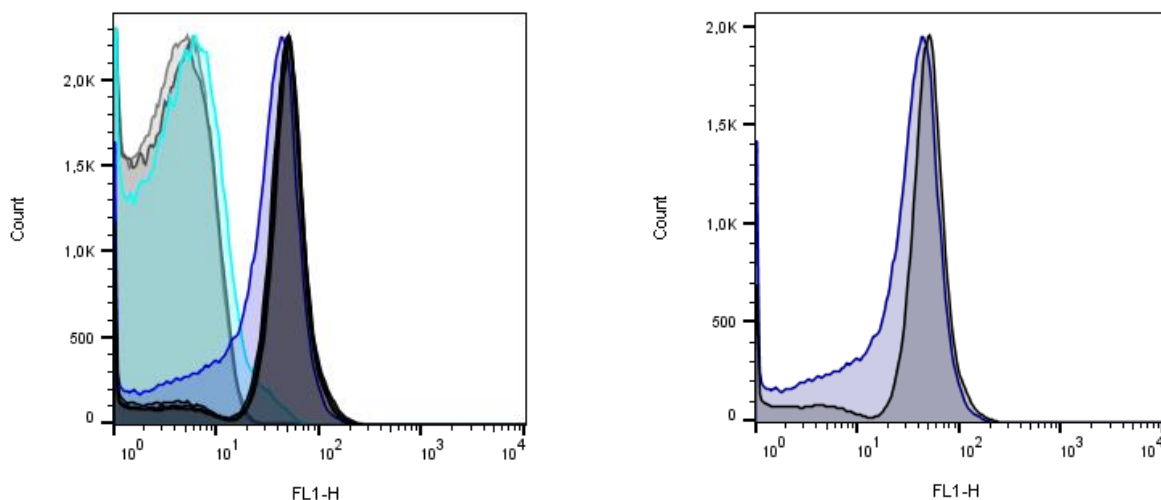
### 7.3.2 The level of population heterogeneity was constant when the *LacZ* promoter was fully induced

The bistable nature of the *LacZ* promoter in *E. coli lacY*<sup>+</sup> strains have been elucidated in recently papers by Choi et al., using a lacY-YFP fusion (Choi *et al.*, 2008). Two phenotypes (induced cells with highly fluorescent membranes and un-induced cells with a small number of membrane-bound permeases) were observed when TMG (lactose analog methyl- $\beta$ -D-thiogalactoside) was present at levels which did not fully induces the promoter, but at a concentration which was just high enough to relieve LacI from the promoter.

The MG1655 *lacZ::gfp* reporter strain is *lacY* and the induction of the promoter was expected to happen gradually and thus not result in two phenotypes at intermediate IPTG concentrations. Wanting to elucidate if the concentration of IPTG, and thus the level of GFP expression, would affect the level of phenotypic variation single cell analysis of the MG1655 *lacZ::gfp* reporter strain was performed.



**Figure 29** Contour plots showing forward scatter and fluorescence for *E. coli* MG1655 *lacZ::gfp* strain induced with different concentrations of IPTG. The samples shown in B-H were taken after the cells had been growing exponentially in chemically defined medium with IPTG for 640 minutes. (A) MG1655 *lacZ::gfp* propagated as a balanced growing culture with 0mM IPTG (B) MG1655 *lacZ::gfp* propageated in medium with 0.01 mM IPTG (C) 0.025 mM IPTG (D) 0.050 mM IPTG (E) 0.075 mM IPTG (F) 0.1 mM IPTG (G) 1 mM IPTG and (H) the control strain 2670 propagated in medium with 1 mM IPTG.

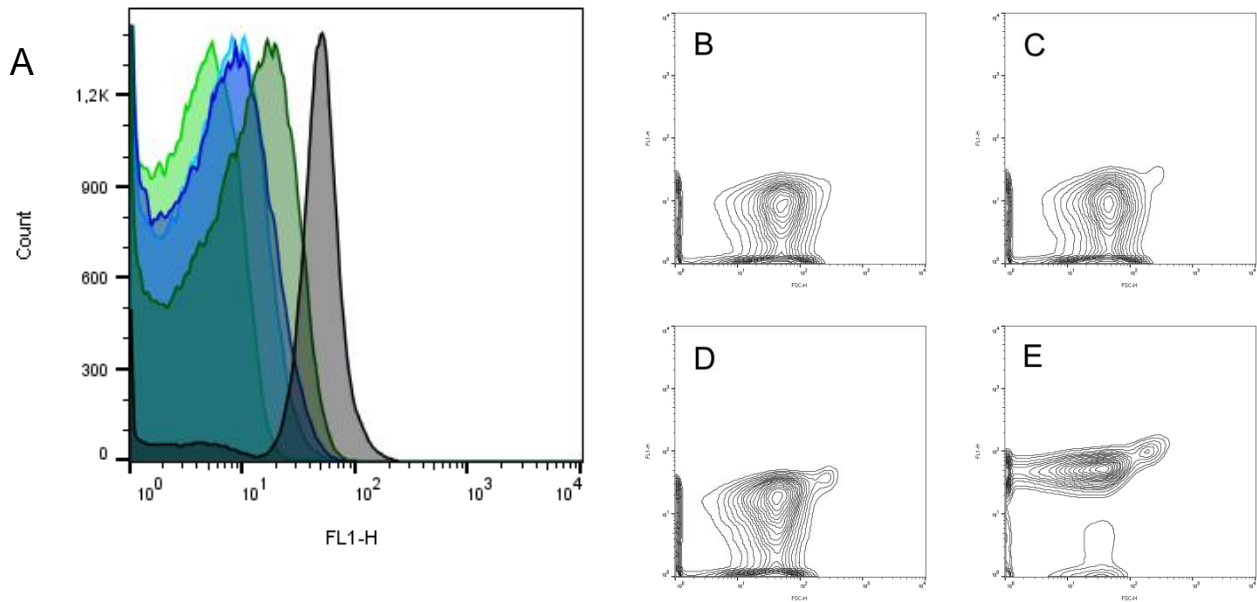


**Figure 30 (Left)** Overlay histograms showing the distribution of fluorescence per cell [FL1-H] for MG1655 *lacZ::gfp* at different IPTG concentrations. The samples are the same as shown in the contour plots in Figure 29. The *E. coli* GFP strain was propagated in medium with different concentrations of IPTG. 0.010 mM IPTG (cyan histogram) and 0.025 mM IPTG (blue histogram). Histograms for the cultures induced with 0.050 mM IPTG, 0.075 mM IPTG, 0.1 mM IPTG and 1 mM IPTG are convergent and are shown as a dark grey histogram with a distinct peak. MG1655 *lacZ::gfp* cultivated in medium without IPTG and 2670 cultivated in the presence of 1 mM IPTG were included as controls (grey histograms).

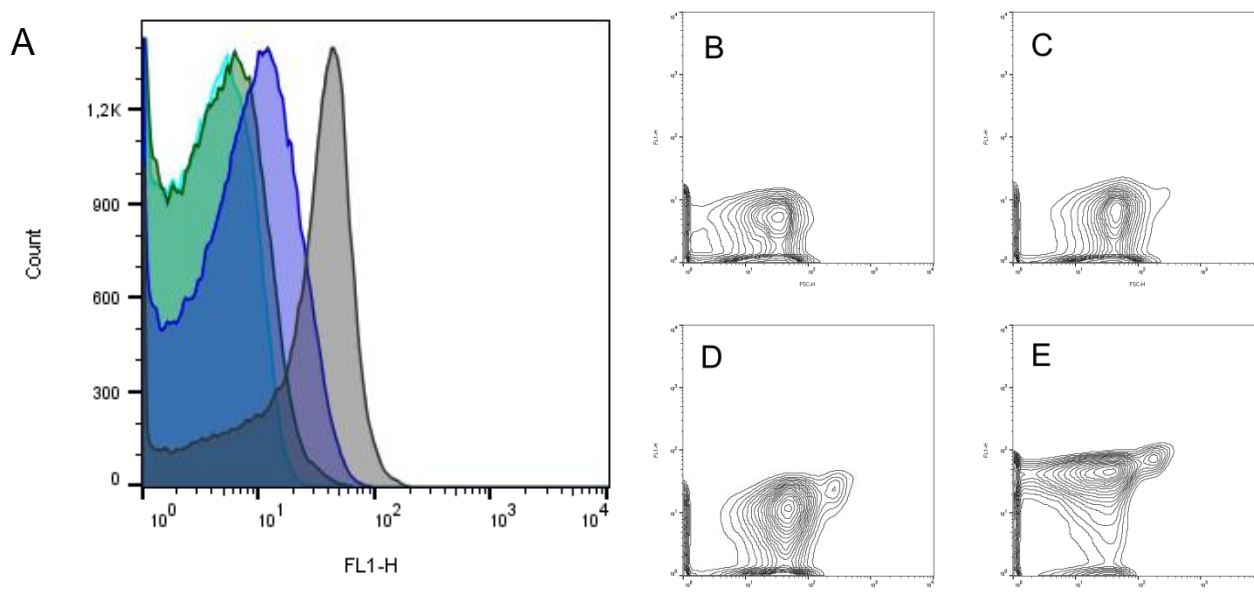
**(Right)** Overlay histograms showing the distribution in fluorescence per cell in MG1655 *lacZ::gfp* induced with 0.025 mM IPTG (blue histogram) and 1 mM IPTG (black histogram) – same histograms as shown in the picture to the left.

The results from flow cytometry analysis of the MG1655 *lacZ::gfp* strain cultivated as exponentially growing culture at various IPTG concentrations can be seen as histograms Figure 30. The samples were taken from exponentially growing cultures, which had been induced with IPTG for more than 640 minutes. The doubling time for MG1655 *lacZ::gfp* in AB medium supplemented with glycerol was measured to be 132 minutes  $\pm$  5.4 minutes. The shapes of the histograms imply that the phenotypic variation was slightly larger in the culture propagated with 0.025 mM IPTG in the medium. The slightly larger variation at this intermediate IPTG was caused by a part of the population being less fluorescent, compared to the cells in the fully induced cultures (see also contour plot Figure 29C). These less fluorescent cells increased the variation in cellular fluorescence at this intermediate IPTG concentration, compared to the variation in cellular fluorescence observed at higher IPTG concentrations (Figure 30, right). The histograms showing the distribution of fluorescence per cell in cultures with fully induced *lacZ* promoters, consists of a distinct peak and a smaller peak (Figure 30, left, black histogram). The events with a relative fluorescence between 1 and 10 comprised the smaller peak and this tiny subpopulation could consist of non-viable cells with permeable cell membrane or cells with repressed *lacZ* promoter.

Wanting to visualize, on a single cell level, the GFP expression in the cells after induction with IPTG, samples were taken from exponentially growing cultures within the first hour after IPTG was added to the medium. Two exponentially growing *E. coli* MG1655 *lacZ::gfp* cultures were induced with 0.025 mM and 1 mM IPTG respectively. These IPTG concentrations were chosen, to observe the distribution of cellular fluorescence in a population with fully induced *lacZ* promoter and in a population with intermediate GFP expression. The results are shown in Figure 31 and Figure 32.



**Figure 31** (A) Overlay histograms showing the fluorescence at a single cell level measured for MG1655 *lacZ::gfp* strain induced with 1 mM IPTG. The grey histogram shows the fluorescence per cell for the culture which had been induced with IPTG for almost five generations (same sample as shown in Figure 30 and in contour plot E). The green histogram shows the fluorescence per cell for a culture of MG1655 *lacZ::gfp* with un-induced *lacZ* promoter. The cyan, blue and dark green histograms show the fluorescence per cell measured for an exponentially growing culture 10 (cyan histogram), 25 minutes (blue histogram) and 40 minutes (dark green histogram) after induction of the *lacZ* promoter. The cyan and blue histograms are convergent. The samples shown in the overlay histogram is also shown as contour plots. (B) the sample taken 10 minutes af IPTG induction (C) 25 minutes and (D) 40 minutes after IPTG induction. (E) This samples was taken after the MG1655 *lacZ::gfp* reporter strain had been cultured in medium with 1 mM IPTG for almost five generations.



**Figure 32 (A)** Overlay histograms showing the fluorescence at a single cell level measured for MG1655 *lacZ::gfp* strain induced with 0.025 mM IPTG. The grey histogram shows the fluorescence per cell for the culture which had been induced with IPTG for more five generations (same as shown in Figure 30). The cyan histogram shows the fluorescence per cell for a culture of MG1655 *lacZ::gfp* without induction of the *lacZ* promoter. The dark green and blue histograms show the fluorescence per cell measured for an exponentially growing culture 10 minutes (dark green histogram) and 60 minutes (blue histogram) after induction of the *lacZ* promoter. Contour plot (B) shows the population profile for the GFP reporter strain with an un-induced *lacZ* promoter. Contour plot (C) and (D) shows the population profile for MG1655 *lacZ::gfp* 10 minutes and 60 minutes after induction of the *lacZ* promoter (same samples as shown in the overlay histogram). Contour plot (E) shows the population profile for a MG1655 *lacZ::gfp* culture which had been induced with IPTG for more five generations (same as shown in Figure 29C).

The fluorescence of the reporter strain was only strong enough to be separated from the fluorescence level of the control strain when the *lacZ* promoter had been fully induced for more than an hour (Figure 31). This precludes analysis beyond the visual display in the histograms, as histogram gating of the events of interest will be difficult.

IPTG was added to balanced growing reporter strain cultures and it was expected, based on experience from  $\beta$ -galactosidase assay that the activation of the *lacZ* promoter and transcription of the reporter gene would happen within 2-3 minutes (Kaempfer & Magasanik, 1967). Folding and maturation of GFPmut3 (the GFP variant expressed in MG1655 *lacZ::gfp*) takes, according to Hebisch et al. between 5 and 6.2 minutes, but depends on the *E. coli* strain used (Hebisch et al., 2013). The doubling time for MG1655 *lacZ::gfp* in AB medium supplemented with glycerol was 132 minutes  $\pm$  5.4 minutes; hence the fluorescence of a cell could potentially continuously increase in 132 minutes after induction with IPTG, before it, due to cell division, would be halved. Comparing the fluorescence of the cells with had been induced with IPTG for more than five

generations with the fluorescence of the cells which had been induced with IPTG in 10, 25 and 60 minutes it was observed that the fluorescence per cell gradually increases during the first 60 minutes after induction and that the cells after having been induced with IPTG for more than one generation, not surprisingly, are substantial more fluorescent.

The flow cytometry results presented here are based on preliminary experiments. Optimization of the flow cytometry settings could improve the possibility to gate the events of interests, making it possible to analyze the population heterogeneity using descriptive statistics. A brighter fluorescent reporter strain could facilitate gating of the GFP expressing cells.



## 8 Stationary phase stress and NaCl addition induces homogeneity in resistance profiles from heterogeneous *Staphylococcus* cultures.

Anne Egholm Pedersen<sup>1</sup>, Kristoffer T. Bæk<sup>2</sup>, Jan Martinussen<sup>1</sup>, Dorte Frees<sup>2</sup> & Mogens Kilstруп<sup>1</sup>

<sup>1</sup>: Metabolic Signaling and Regulation Group, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>2</sup>: Department of Veterinary Disease Biology, University of Copenhagen, 1870 Frederiksberg, Denmark

Keywords: MRSA, population profile analysis (PAP), balanced growing cultures, subpopulations

### 8.1 Introduction

*S. aureus* is a major cause of hospital- and community-acquired infections worldwide (Lim & Strynadka, 2002). Community-associated methicillin-resistant *S. aureus* (CA-MRSA) strains are causing a severe pandemic of mainly skin and soft tissue and occasionally fatal infections (Otto, 2010). Methicillin-resistant *S. aureus* (MRSA) (also called oxacillin-resistant *S. aureus* (Mcdougal *et al.*, 2003)) is any strain of *S. aureus* that has developed resistance to  $\beta$ -lactam antibiotics. *S. aureus* strains unable to resist these antibiotics are classified as methicillin-sensitive *S. aureus*, or MSSA.  $\beta$ -lactam antibiotics inhibit the native penicillin-binding proteins (PBPs) which is involved with the synthesis of staphylococcal cell wall peptidoglycan. All MRSA strains have acquired the *mecA* or *mecC* gene encoding a peptidoglycan transpeptidase, a low affinity penicillin binding protein PBP2a/PBP2'. PBP2a/PBP2' can continue the catalysis of peptidoglycan transpeptidation in the presence of  $\beta$ -lactam concentrations that inhibit the  $\beta$ -lactam-sensitive PBPs normally produced by *S. aureus* (Bæk *et al.*, 2014; Lim & Strynadka, 2002; Paterson *et al.*, 2014).

Resistance to the  $\beta$ -lactam antibiotics is, for many clinical isolates of MRSA, expressed in a heterogeneous fashion. A heterologous profile of a MRSA population, grown from single cell inoculum, could either be caused by a large standard deviation in the resistance level or by the presence of subcultures with different resistance levels. Population analysis profile (PAP) is an analysis method which can be used to investigate phenotypic variation and visualize the presence of subcultures with different levels of antibiotic resistance. MRSA strains can have either a homogenous or a heterologous resistance profile. Some strains which consist of bacteria with a uniform and high resistance towards methicillin have a homogenous profiles, while other strains have heterologous profiles (Hartman & Tomasz, 1986; Tomasz *et al.*, 1991). Studies have shown that quantitative analysis of hetero-resistant MRSA cultures can provide profiles that are characteristics of that particular MRSA strain (Figueiredo *et al.*, 1991; Tomasz *et al.*, 1991).

Population analysis profile (PAP) is a method for quantitative analysis which is used to investigate the antibiotic resistance pattern. Plotting of colony counts (CFU/ml) against the antibiotic concentration in the test agar plates provides a graphic display (PAP) of the composition of resistance in the bacterial culture. Quantitative analysis of MRSA by PAP is usually tested on stationary phase overnight cultures. This may lead to over-estimation of the resistance since a culture of isogenic bacteria cells that enters stationary phase after having been growing exponentially will usually become more resistant to heat, disinfectants and osmotic stress, and will experience morphological and physiological changes; cell volume decrease, cell shape change, nucleoid compaction and alteration in cell wall composition (Ishihama, 1997, 1999; Kolter *et al.*, 1993; Neidhardt *et al.*, 1990; Roszak & Colwell, 1987).

When MRSA strains are analyzed on media containing  $\beta$ -lactam antibiotics, the addition of NaCl to the medium has been proposed as a means to increase the reproducibility of the analysis (Chambers & Hackbarth, 1987; Ming Bo Huang *et al.*, 1993). The increased reproducibility is followed by increased resistance towards the  $\beta$ -lactam antibiotic which could be the result of NaCl induced stress to the cells. It has been shown that the profile of a MRSA culture can be influenced by stress through lowering the incubation temperature. In 1986 it was shown that the phenotypic expression of  $\beta$ -lactam resistance of a hetero-resistant strain is conditional on the plate incubation temperature when performing quantitative analysis (Hartman & Tomasz, 1986). Hartman and Tomasz showed that the hetero-resistant profile of the MRSA strain DU could be changed to a homogenous profile, by changing the incubation temperature from 37°C to 30°C (Hartman & Tomasz, 1986).

We wanted to address the question whether the quantitative analysis profile (PAP) of the highly virulent CA-MRSA strain USA300 would be affected if the analysis was performed with a balanced exponentially growing culture as inoculum instead of a stationary phase culture. Subsequently the effect of NaCl addition was analyzed under both conditions. We found that the use of balanced cultures for inoculum resulted in two populations with different resistance levels on test plates without NaCl, and that the use of stationary cultures as inoculum, or the addition of NaCl to the test plates, resulted in more homogenous profiles and higher resistance.

## **8.2 Materials and methods**

### **8.2.1 Strain**

*S. aureus* strain JE2. JE2 is the CA-MRSA strain USA300 cured of plasmids (Fey *et al.*, 2013).

### **8.2.2 Media and growth conditions**

Mueller Hinton Broth (MHB) was used for propagation of JE2. BBL™ Mueller Hinton II Broth was used for agar plates (MHA). The MHA was made with and without the addition of 2% NaCl.

Balanced growth was obtained by inoculating a single colony of JE2 in pre-warmed MHB. The colony was propagated at 37°C with aeration until OD<sub>600</sub> was approximately 0.2. This culture was used as inoculum and 25 ml pre-warmed MHB was inoculated with this exponentially growing pre-culture at an OD<sub>600</sub> of 0.02. When the optical density for the exponentially growing culture was between 0.3 and 0.4 the cells were spun down (10 minutes at 4000 rpm at room temperature) and resuspended in 0.9 % NaCl to an OD<sub>600</sub> of 1.

The stationary phase culture was made by inoculating 1 mL of MHB with a single JE2 colony. The colony was propagated at 37°C with aeration until the day after.

### **8.2.3 Population analysis profiles**

Antibiotic resistance profiles were determined by plating appropriate dilutions of either a balanced growing culture or a stationary phase culture on Müller-Hinton agar (MHA) plates containing increasing concentrations of oxacillin (from Sigma). Plates were incubated at 37°C for 48 h, and the number of bacterial colonies was counted. Plotting colony counts against drug concentrations provides a graphic display (PAP) of the composition of the bacterial culture as to the homogeneity or heterogeneity of the antibiotic susceptibility phenotype.

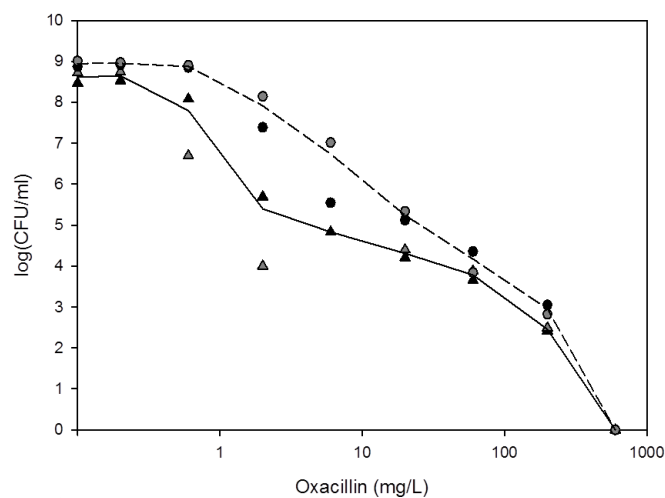
### **8.2.4 Agar diffusion spot test**

MHA Plates with increasing concentration of oxacillin were used for spot test. Seven dilutions ( $10^{-1}$  to  $10^{-7}$ ) of the balanced growing JE2 culture, and seven dilutions ( $10^{-1}$  to  $10^{-7}$ ) of the stationary phase JE2 culture, were spotted. The cultures were diluted in 0.9% NaCl dilutions. The plates were incubated at 37°C for 48 h, and the number of bacterial colonies, in the spots made with a dilution thin enough to produce single colonies, was counted. The results obtained from the agar diffusions spot test was used to determine the dilutions used for the quantitative analysis used to produce PAPs.

## 8.3 Results

### 8.3.1 Induction of a homogenous profile by stationary phase stress in *S. aureus* JE2 cultures

To investigate if the JE2 population analysis profile would be affected by the state of the culture used for the analysis, samples from balanced exponentially growing cultures or stationary phase cultures of JE2 were diluted and plated on MH plates with different concentrations of oxacillin. The procedure for the quantitative analysis based on agar dilution test is described in material and methods (section 8.2.3). The colony count at each concentration of oxacillin was plotted as a function of oxacillin concentration for the two conditions (see Figure 33).



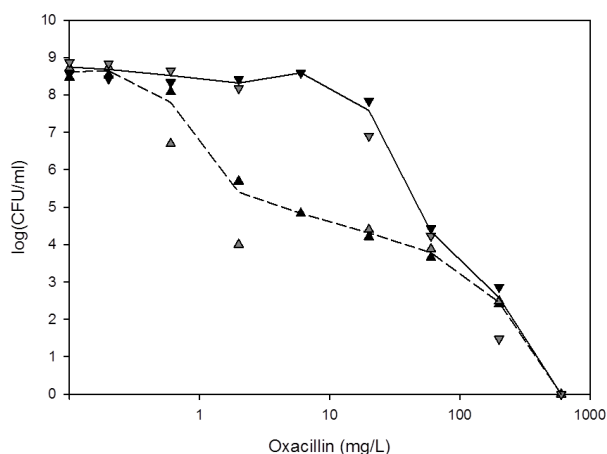
**Figure 33 Population analysis profiles made base on quantitative analysis of balanced growing JE2 cultures and stationary phase JE2 cultures. The lines represent the average calculated from experimental duplicates. Dashed line and circles: PAP base on quantitative analysis with stationary phase cultures of JE2. Solid line and triangles: PAP base on quantitative analysis with balanced growing cultures of JE2.**

It is clear that the antibiotic resistance profile in the balanced growing cultures was heterogeneous, while a more homogenous profile at a higher resistance level was induced as the culture entered the stationary phase.

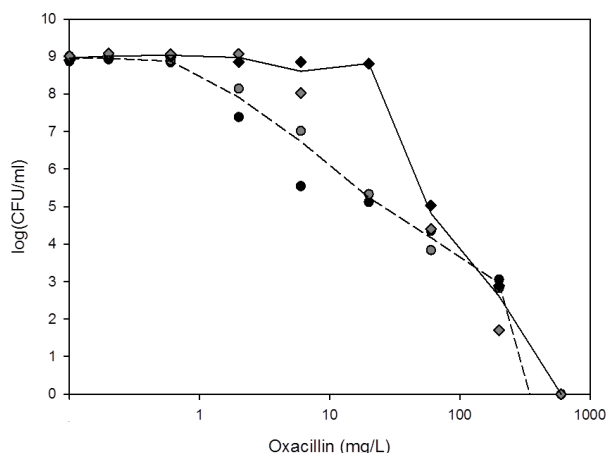
### 8.3.2 NaCl induces a homogenous profile in balanced JE2 cultures

To see if osmotic stress could induce the homogenous and high resistance phenotype like stationary phase stress, samples from a balanced exponentially growing culture were diluted and plated on MH oxacillin plates in the presence of NaCl, as recommended to enhance detection and improve the accuracy of the test (Chambers & Hackbarth, 1987; Ming Bo Huang *et al.*, 1993). It is clear from the profile of the PAP analysis shown in Figure 34 that NaCl addition to the test plates induced a homogenous resistance profile in the

balanced culture. When a stationary culture was diluted and plated on MH oxacillin plates, the additional oxidative stress imposed by including NaCl in the test plates resulted in a slightly more homogenous profile (Figure 35).



**Figure 34 Population analysis profiles base on quantitative analysis with balanced growing JE2. Each experiment was performed in duplicate. The solid line represents the average obtained for quantitative analysis made with MHA containing 2 % NaCl (triangle down). The dashed line represents the average obtained for quantitative analysis performed with MHA without NaCl (triangle up).**

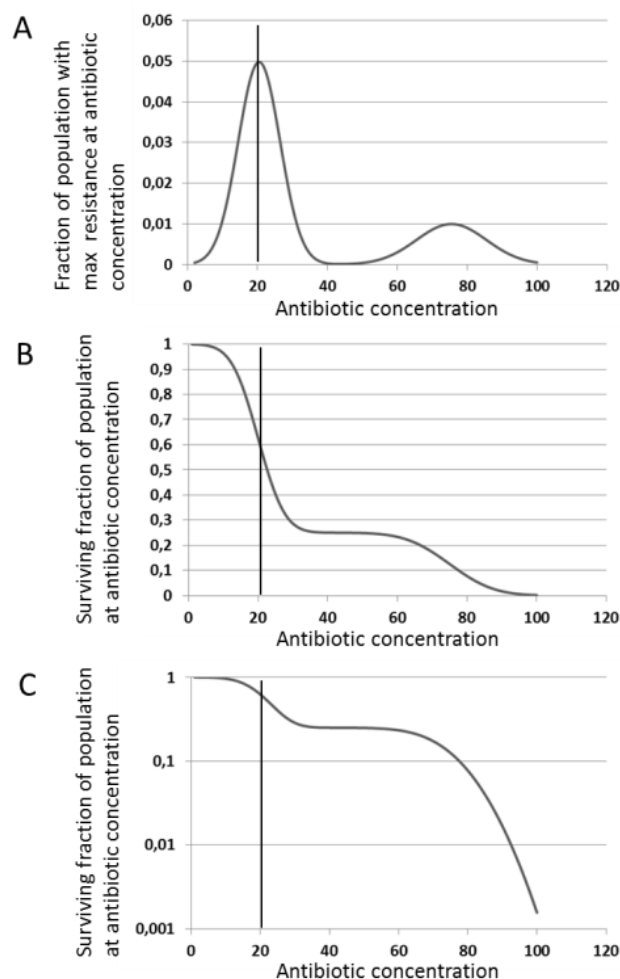


**Figure 35 Population analysis profiles base on quantitative analysis with a stationary phase JE2 culture. Each experiment was performed in duplicate. The solid line represents the average obtained for quantitative analysis made with MHA containing 2 % NaCl (diamonds). The dashed line represents the average obtained for the quantitative analysis performed with MHA without NaCl (circles).**

### 8.3.3 The slope of PAP curves reveals underlying populations

Population analysis profiles are usually interpreted as representing the differential resistance of subpopulations, but a mathematical explanation has never been presented for this assumption. In the following we will show that by plotting the slope of a PAP curve for each concentration you obtain the underlying resistance distribution for each subpopulation in the culture. As an example, Figure 36A shows an idealized antibiotic resistance profile of a bacterial population. The fraction of bacteria that will survive a certain concentration (shown as a vertical line) can be calculated as the area under the curve for the antibiotic concentrations higher than the concentration marked by the vertical line. Mathematically the area can be found by integrating the curve from that certain concentration to infinity. This surviving fraction is exactly what a PAP curve shows. The opposite must also be true, that differentiation of a PAP curve will yield the resistance distribution. This is equal to a diagram showing the slope of the PAP curve for each concentration. The resistance distribution curve in Figure 36A, which is plotted as the slopes of the PAP curve from Figure 36B for each concentration, shows clear correlation to the PAP curve. However when the

PAP curve is plotted in a semi logarithmic diagram (Figure 36C) like usual PAP curves, the visual correlation to the resistance distribution becomes obscured.



**Figure 36 Correspondence between resistance distribution profiles and PAP curves. (A) Hypothetical resistance profile for a bacterial population. The resistance profile shows the fraction of cells (vertical axis), which has an upper limit of resistance that matches the antibiotic concentration on the horizontal axis. (B and C) PAP curve showing the fraction of surviving cells as a function of the antibiotic concentration. The survival fraction is shown in a linear (B) and a logarithmic (C) plot.**

## 8.4 Discussion

The results of the quantitative analysis of JE2 selected on MHA with NaCl in Figure 34 and Figure 35 shows that regardless of whether the analysis was made with a balanced growing culture or with a stationary phase culture did the profile of JE2 resembling a homologous profile. The profiles obtained using MHA with NaCl were not noticeable affected by the state of the culture. This was surprisingly since it was not expected that the bacteria cells in a balanced growing culture would exhibit the same extent of resistance toward the  $\beta$ -lactam antibiotic as the stationary phase culture. The stationary phase culture was expected to be more

resistant towards the antibiotic since stationary phase cultures have proven to be more resistant towards different types of stress, (Watson *et al.*, 1998).

Before the balanced growing cultures were plated on MHA were the MHB used for propagation removed and the cells were resuspended in 0.9 % NaCl solution and dilution series were made. The dilution series processing must have interrupted the balanced state of the culture, resulting in quantitative analysis profiles resemble those observed for the stationary phase JE2 cultures. It could be that that balanced growing cultures during the time they were centrifuged and resuspended in 0.9% NaCl solution did express gene which made the cells more resistant and since the cells upon plating had time to adapt to the new environment it seem to become subordinate how the culture used for the quantitative analysis had been propagated.

In strains of *S. aureus* with heterologous profiles cells with low resistance towards  $\beta$ -lactam comprise a larger part of the population than in strains with homologous profiles, it was thus purposed that NaCl protect susceptible cell from autolysis, making the population develop a more homogenous in profile (Chambers & Hackbarth, 1987). Cell wall-active antibiotics trigger autolysis of susceptible bacteria, including *S. aureus* (Tomasz & Waks, 1975). Autolysis of susceptible bacteria has been found to decrease in the presence of NaCl (Chambers & Hackbarth, 1987). Besides observing that heterogeneous strains differed in their pattern of lysis from homogeneous strains Chambers and Hackbarth also observed that the amount of low affinity peptidoglycan transpeptidase PBP2a produced could not account for the heterologous profiles.

The results obtained from the spot test (data not shown) and the quantitative agar analysis presented here, correlates well with literature, as JE2 clearly exhibit a more resistant profile when sodium chloride was present in the medium (Figure 34 and Figure 35).

When sodium chloride was not added to the MHA the JE2 strain became more susceptible towards oxacillin and the population analysis profile resembles more a heterogeneous profile, than the profile obtained when the quantitative analysis was made with MHA with sodium chloride.

As both homogenous and heterogeneous profile MRSA strains are expressing PBP2a/PBP2' other factors must influence the antibiotic resistance, and hence the profile of the strains. It has been shown that if RelA is expressed constitutively in a MRSA strain with a heterologous profile (expressing a plasmid borne *mecA*), the profile of this strain will change and become homogenous (Kim *et al.*, 2013; Mwangi *et al.*, 2013). It has also been show that inactivation of ClpXP proteases increases the resistance towards  $\beta$ -lactam antibiotics, changing JE2 from having a heterogeneous profile (when tested on tryptic soy agar plates) into a homogeneously resistant strain (Bæk *et al.*, 2014).

Based on the results presented here it can be concluded that the state of the culture used for the phenotypic analysis had larger impact on the population profile when NaCl was not added to the medium.

## 9 Discussion and concluding remarks

The objective of this project was to elucidate heterogeneity in isogenic populations of microorganisms. The approach used in part of this work was based on an idea, which can be used to elucidate how the design of fermentation processes affects the production yield. The idea behind this project was to analyze heterogeneity in an isogenic population of reporter cells expressing two fluorescent reporter proteins. The expression of the reporter proteins should be regulated by an industrial relevant promoter and a growth rate regulated promoter respectively. The analysis of such a dual reporter strain could elucidate if the fastest growing cells are the most productive or if the most productive cells are those with a lower growth rate than the average. Additionally could the analysis elucidate how the experimental design might affect this ratio.

A dual fluorescent reporter strain was never constructed. However, the direction for the work was established by this idea, and the path leading in that direction turned out to contain many interesting investigations, experiments and results.

Initially two reporter strains were constructed and validated. The results obtained showed that the *B. subtilis* and *L. lactis* reporter strains expressed GFP on a detectable level, and that the expression was growth rate regulated. Flow cytometry analysis of the reporter strains was used to study variability in single cell growth rates. Focusing on how the choice of medium and the history of the pre-culture affected the variability in single cell growth rates it became possible to elucidate how the guidelines of the Copenhagen School of Bacterial Growth Physiology and thus the use of balanced growing cultures, affect the degree of population heterogeneity.

During investigations on how experimental conditions affect population heterogeneity some of the challenges using GFP as a reporter became apparent. The stability of GFP made it impossible to detect decreases in growth rate directly. Small fluctuations in the systems were probably also masked by the stability of the protein. The time delay between transcription initiation and emission of fluorescence, primarily caused by the time it takes for the GFP chromophore to mature, did also mask for detection of smaller fluctuations caused by the phenotypic noise in the cell. The fluctuation in gene expression and maturation of the fluorescent protein itself could on the other hand have affected the phenotypic noise. Using GFP to detect fluctuations in phenotypic characteristics in conditions, which were altered during the experiment e.g. changes in incubation temperature or pH would quench the GFP signal either by denaturation or by affecting the protonated state of the chromophore (Campbell & Choy, 2001; Smith, 2002; Tsien, 1998). If a change in cell-to-cell variability was detected after an external variable had been altered, it would be difficult to determine if the change was caused by the effect the alteration had on the phenotypic noise or if the change was caused by the effect the alteration had on the fluorescent protein.



The growth rate regulated expression of GFP in *B. subtilis* AE099 and *L. lactis* AE072 were validated on a population level and it was shown that GFP expression correlated with the increase in biomass and that the specific fluorescence (fluorescence per OD) was growth rate regulated. The growth rate regulated GFP expression in the reporter strains was used as a measure for single cell growth rate. Thus investigation in cell-to-cell variability in an isogenic population of microorganisms could be done using this variable. When the reporter strain was validation on a single cell level it was considered whether to use the total cellular fluorescence to analyze variation in single cell growth rate, or if the cell size should be included in the interpretation of the results. It was chosen to include the cell size in the interpretation, as the total cellular fluorescence detected by flow cytometry would depend on the cell size since cells are autofluorescence. This could be determined by measurements of the control strain. For the *B. subtilis* control strain AE044 it was observed that the cells growing exponentially with a doubling time of 24 minutes were (on average) both larger and more fluorescent than cells growing with a doubling time of 73 minutes (data not shown). However, the main argument for including the cell size in the interpretation of results was that the activity of the ribosomal promoters (and thus the level of GFP expression) in cells growing exponentially, and at the same rate, was similar in an elongated cell which prepared for division and in a smaller cell just gone through a cell division.

In a recent paper by Paulander et al. flow cytometry results showed that *E. coli* cells treated with bactericidal antibiotics increased in size and fluorescence. In one of the reporter strains investigated in this paper the fluorescence was generated from a chromosomally integrated GFP marker. In the other strains investigated the cellular fluorescence was generated by oxidation of the hydroxyphenyl fluorescein (HPF) probe detecting the concentration of reactive oxygen species (ROS) in the cell (Paulander et al., 2014). Paulander et al. also included a temperature sensitive cell division mutant, which increased in cell size and fluorescence in the absence of antibiotic and without addition of HPF. In 2007 Kohanski et al. published a paper in which they conclude that reactive oxygen molecules (detected by HPF) may play a role in the killing effect of bactericidal antibiotics. In this paper flow cytometry results showed that the total cellular fluorescence was higher in cells treated with bactericidal antibiotics than in cells, which had not been treated with a drug (Kohanski et al., 2007). Kohanski et al. did not include the cell size in their investigations. Paulander et al. showed that the antibiotic treated cells increased in size and that the elevated HPF signal measured in the antibiotic treated cell primarily stems from the larger cell size rather than the elevated level of ROS-mediated HPF oxidation (Paulander et al., 2014).

The *B. subtilis* and *L. lactis* reporter strains were analyzed at different growth conditions to elucidate the effect the choice of medium and the physiological state of the cells in the pre-culture have on population heterogeneity. According to the Copenhagen School of Bacterial Growth Physiology the use of balanced growth offers: “the best possible conditions for making reproducible experiments with bacterial cultures” (Ingraham et al., 1983). As shown by the results presented in the manuscript in section 3. The lowest level of

population heterogeneity was observed in a *B. subtilis* culture propagated in agreement with the guidelines of the Copenhagen School of Bacterial Growth Physiology.

An interesting observation was that the level of population heterogeneity in a culture inoculated with a stationary phase pre-culture first came close to the level observed in the balanced growing culture after several generations in exponential phase. This result indicates why the use of a balanced growth offers the best possible conditions for making reproducible experiments.

The experiments, which were not replicated should be viewed as a snapshots and extrapolation of data should be done with precaution. Consistent tendencies observed in several of the experiments thus clearly indicate that the data were useful and the results probably could be replicated if time was available. This statement is supported by the flow cytometry analysis of the balanced growing *B. subtilis* cultures, since this growth experiment was performed additional times, and similar results were obtained.

Preparing a balanced growing culture requires care, especially during the inoculation. If the cells were exposed to a brief temperature shift during the inoculation process the balanced state of growth would be affected and an increase in cell-to-cell variability was observed (data not shown). When the cells are in a balanced state of growth changes in temperature and in aeration will, as well as depletion of a medium component, disrupt this state. An issue when propagating the *B. subtilis* cells in chemically defined medium containing thiamine and tryptophan, and in one case also glutamate, was that one of the components potential could be depleted. The depletion would probably first take place at higher cell densities, but verification by HPLC analysis could have determined when the balanced growth of the culture was affected.

Samples were prepared for flow cytometry analysis by resuspending the cells in PBS solution. Handling the samples probably affected the level of phenotypic noise, even though the samples were kept on ice and centrifuged at 4°C. It could be interesting to repeat some of the experiment using a lab-scale fermenter and make the analysis using an on-line flow cytometer. The on-line sampling would make it possible to conduct a more continuous analysis and at the same time eliminate the noise introduced by handling the samples. The lab-scale fermenter offers the possibility to cultivate the cells in a more controlled environment. On-line flow cytometry and lab-scale fermenters have already been used to analyze fluorescent reporter strains. However, the publications that I have knowledge about, do not include results showing quantification of the population heterogeneity. Han et al. have used on-line flow cytometry to study *E. coli* GFP reporter strains under different operating conditions in well mixed and scale down bio-reactors (Han *et al.*, 2013). In this paper they focus on the experimental condition and they use the mean GFP fluorescence to show how glucose perturbations during the cultivation affect the growth regulated expression of GFP. Population heterogeneity was not the focus of this paper, but the experimental setup could potentially have been used to elucidate how the glucose perturbation affected the sample distribution and the population heterogeneity. In a paper by Delvigne et al. on-line flow cytometry was used to study an *S. cerevisiae* GFP reporter strain dyed with PI (Delvigne *et al.*, 2014). Delvigne et al. used box-plot and mean to median ration to visualize population

heterogeneity in the GFP expressing *S. cerevisiae* strain FE440. They observed two peaks for the PI strained cells but only one peak for the GFP expression, and concludes that more understanding concerning the reporter strain is needed before it is decided if quantitative analysis can be done using mathematical tools or if more “sophisticated computational tools are needed for the quantification” (Delvigne *et al.*, 2014).

Another approach to study population heterogeneity is to analyze phenotypic noise in protein expression. de Jong *et al.* used GFP expression under regulation of various promoters to study how phenotypic heterogeneity in a *B. subtilis* culture is affected by carbon source starvation (de Jong *et al.*, 2012). de Jong *et al.* used time series fluorescent microscopy to detect the cellular fluorescence. This data set was used to calculate phenotypic noise strength (PNS) (Ozbudak *et al.*, 2002) and compare the PNS for different promoters in different growth phases. Phenotypic noise in protein expression was also elucidated in a paper by Elowitz *et al.* Results presented in this paper show the degree of intrinsic and extrinsic noise in *E. coli* cells. The noise level was determined by measurements of two fluorescent proteins, regulated by two identical promoters, by fluorescent microscopy (Elowitz *et al.*, 2002).

However, only a limited number of publications have to my knowledge focused on quantification of heterogeneity in isogenic population of microorganisms using flow cytometry analysis. Julia *et al.* have analyzed microbial cultures with flow cytometry and by fitting different distribution types to the data search for a way to compare the data using a statistical method (Julià & Vives-Rego, 2005; Julià *et al.*, 2010; Vives-Rego *et al.*, 2003). These papers do not compare the level of heterogeneity but focus on finding a distribution, which can describe various samples.

The approach used in this work to investigate population heterogeneity made it possible to quantify the heterogeneity within a sample and to compare the degree of heterogeneity in various samples. The choice to use descriptive statistics was based on thorough assessment of published material (Thyregod *et al.*, 2015). In addition to the use of descriptive statistics the software JMP® can be used to quantify the size of eventual subpopulations and fit a distribution to the sample.

In a paper from 2009 visual assessment of population profiles shown in pseudocolor plots was used to show GFP expression in a prokaryote reporter strain. The GFP expression in this reporter strain was transcriptional regulated by a methanol induced promoter (Strovas & Lidstrom, 2009). During shifts from methanol to succinate or vice versa Strovas and Lidstrom observed that a part of the population did not respond to the shift and concluded that these cells did not divide. Using microscopy to analyze cells attached to a glass slide they observed that some cell did not divide, some did, but then stopped and others continuously divided during the time of the experiment. Strovas and Lidstrom did not quantify the level of heterogeneity observed in the flow cytometry samples, as have been done in this work, but they evade one of the main drawbacks using flow cytometry to study single cells. The drawback is that the data acquired represent only a momentary snapshot of the cells. Thus it could have been interesting to combine the flow cytometry analysis shown in this work with analysis using time lapse fluorescent microscopy.

If the time had been available it would have been interesting to study how artificial introduced gradients affect the population heterogeneity. Most optimal could the study of gradient zone be combined with automated sampling by an on-line flow cytometer and a reporter strain expressing an unstable variant of GFP. Unstable variants of GFP can be made by adding a peptide tail to the protein, making it a target for intracellular proteases (Andersen *et al.*, 1998; Keiler & Sauer, 1996). The reporter strains presented in this work are all expressing a stable variant of GFP. As GFP is stable the history of the cell lineage would to some extent be included in the cellular fluorescence and hence the quantitative determination of variation in single cell growth rate would presumably be underestimated. It would thus be interesting to repeat the results presented in section 3 with a *B. subtilis* reporter strain expressing an unstable GFP variant, as this will give a more accurate measure of the dynamic changes.

Using an unstable GFP as reporter could be challenging, as a relative strong signal is needed to detect the cells using a flow cytometer. The choice of using an unstable GFP variant would be a compromise between having a GFP variant with a half-life long enough to make the cell detectable, but still short enough to make it possible to detect the dynamic changes of the system. The signal could be increased by placing several copies of the reporter gene on the chromosome of the reporter strain.

Construction of a reporter strain expressing two fluorescent proteins was not attempted, but analysis of a reporter strain expressing a fluorescent protein under regulation of an inducible promoter was performed. Preliminary validation of phenotypic noise in gene expression was studied by analysis of the *E. coli* reporter strain expressing GFP under regulation of the *lacZ* promoter. This preliminary analysis was made to study which challenges there potentially could be when analyzing GFP expression, regulated by an inducible promoter, on a single cell level by flow cytometry. If the two already constructed reporter strains AE099 and AE072 should be modified to express an additional fluorescent protein the promoters regulating this expression should be relative strong to obtained a detectable fluorescence. Likely candidates could be the *nisA* promoter in *L. lactis* and the *spaS* promoter in *B. subtilis*. The GFP expression regulated by the *spaS* promoter have been analyzed on a single cell level by flow cytometry analysis (Bongers *et al.*, 2005).

During this work different approached have been used to elucidate population heterogeneity. The results showed that heterogeneity in isogenic populations of microorganisms was present and that the degree of heterogeneity could be affected. Single cell analysis by flow cytometry was used to show that the degree of population heterogeneity can be affected by the experimental design. Analysis of the *B. subtilis* GFP reporter strain was used to elucidate how balanced growth lowers the degree of population heterogeneity.

Heterogeneity was addressed on a population level by quantitative analysis of the MRSA strain JE2. The quantitative analysis showed that stress affected the heterogeneous antibiotic resistance profile of this MRSA strain. Even though some issues made it problematic to draw conclusion on the variability in single cell growth rate in the *L. lactis* and *S. cerevisiae* reporter strains the results obtained showed distinct cell-to-cell variability.

This work provides a platform for further analysis of population heterogeneity. Analysis made with descriptive statistics, in combination with distribution analysis, can be used by non-statisticians to quantifying heterogeneity using data obtained with flow cytometry. The *B. subtilis* reporter strain expressing a growth rate regulated GFP can be further modified to become the dual reporter strain mentioned earlier. Analysis of the dual reporter strain expressing two fluorescent reporter proteins regulated by a growth rate and an industrial relevant promoter respectively can thus be done with the same statistical methods as suggested in this work. In the longer run the quantification of heterogeneity in the dual reporter strain can be used to elucidate how a fermentation process can be designed to improve production yield.

## 10 References

- Albano, C. R., Randers-Eichhorn, L., Bentley, W. E. & Rao, G. (1998).** Green fluorescent protein as a real time quantitative reporter of heterologous protein production. *Biotechnol Prog* **14**, 351–4.
- Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjorn, S. P., Givskov, M. & Molin, S. (1998).** New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240–6.
- Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. (2004).** Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625.
- Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O’Shea, E., Pilpel, Y. & Barkai, N. (2006).** Noise in protein expression scales with natural protein abundance. *Nat Genet* **38**, 636–43.
- Beresford, T. & Condon, S. (1993).** Physiological and genetic regulation of rRNA synthesis in *Lactococcus*. *J Gen Microbiol* **139**, 2009–17.
- Billinton, N. & Knight, a W. (2001).** Seeing the wood through the trees: a review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence. *Anal Biochem* **291**, 175–197.
- Boehlke, K. W. & Friesen, J. D. (1975).** Cellular content of ribonucleic acid and protein in *Saccharomyces cerevisiae* as a function of exponential growth rate: calculation of the apparent peptide chain elongation rate. *J Bacteriol* **121**, 429–433.
- Bokman, H. & Ward, W. (1981).** Renaturation of *Aequorea* Green-Flourescent Protein. *Biochem Biophys Res Commun* **101**, 1372–1380.
- Bongers, R. S., Veening, J., Wieringen, M. Van & Kuipers, O. P. (2005).** Development and Characterization of a Subtilin-Regulated Expression System in *Bacillus subtilis* : Strict Control of Gene Expression by Addition of Subtilin. *Appl Environ Microbiol* **71**, 8818–8824.
- Brauer, M. J., Huttenhower, C., Airoidi, E. M., Rosenstein, R., Matese, J. C., Gresham, D., Boer, V. M., Troyanskaya, O. G. & Botstein, D. (2008).** Multiple pathways differentially regulate global oxidative stress responses in fission yeast. *Mol Biol Cell* **19**, 308–317.
- Brøndsted, L. & Hammer, K. (1999).** Use of the integration elements encoded by the temperate lactococcal bacteriophage TP901-1 to obtain chromosomal single-copy transcriptional fusions in *Lactococcus lactis*. *Appl Environ Microbiol* **65**, 752–758.
- Buchanan, R. E. (1918).** Life phases in a bacterial culture. *J Infect Dis* **23**, 109–125.
- Bylund, F., Collet, E., Enfors, S.-O. & Larsson, G. (1998).** Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. *Bioprocess Eng* **18**, 171–180.

- Bæk, K. T., Gründling, A., Mogensen, R. G., Thøgersen, L., Petersen, A. & Paulander, W. (2014).**  $\beta$ -Lactam Resistance in Methicillin-Resistant *Staphylococcus aureus* USA300 Is Increased by Inactivation of the ClpXP Protease. *Antimicrob Agents Chemother* **58**, 4593–4603.
- Campbell, T. & Choy, F. (2001).** The effect of pH on green fluorescent protein: a brief review. *Mol Biol* **2**, 1–4.
- Carlquist, M., Fernandes, R. L., Helmark, S., Heins, A.-L., Lundin, L., Sørensen, S. J., Gernaey, K. V & Lantz, A. E. (2012).** Physiological heterogeneities in microbial populations and implications for physical stress tolerance. *Microb Cell Fact* **11**, 1–13.
- Chambers, H. F. & Hackbarth, C. J. (1987).** Effect of NaCl and nafcillin on penicillin-binding protein 2a and heterogeneous expression of methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **31**, 1982–1988.
- Choi, P. J., Cai, L., Frieda, K. & Xie, S. (2008).** A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. *Science (80- )* **322**, 442–446.
- Christensen, H., Olsen, R. A. & Bakken, L. R. (1995).** Flow cytometric measurements of the cell volumes and DNA contents during culture of indigenous soil bacteria. *Microb Ecol* **29**, 49–62.
- Chung, J. D., Stephanopoulos, G., Ireton, K. & Grossman, a D. (1994).** Gene expression in single cells of *Bacillus subtilis*: evidence that a threshold mechanism controls the initiation of sporulation. *J Bacteriol* **176**, 1977–1984.
- Clark, D. J. & Maaløe, O. (1967).** DNA replication and the division cycle in *Escherichia coli*. *J Mol Biol* **23**, 99–112.
- Cooper, S. (1993).** The origins and meaning of the Schaechter-Maaloe-Kjeldgaard experiments. *J Gen Microbiol* **139**, 1117–1124.
- Cormack, B. P., Bertram, G., Egerton, M., Gow, N. a, Falkow, S. & Brown, a J. (1997).** Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans*. *Microbiology* **143**, 303–311.
- Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996).** FACS-optimized Mutants of Green Fluorescent Protein (GFP). *Gene* **173**, 33–38.
- Craggs, T. D. (2009).** Green fluorescent protein: structure, folding and chromophore maturation. *Chem Soc Rev* **38**, 2865–2875.
- Cronan, J. E. (1968).** Phospholipid Alterations During Growth of *Escherichia coli*. *J Bacteriol* **95**, 2054–2061.
- Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A. & Tsien, R. Y. (1995).** Understanding, improving and using green fluorescent proteins. *TIBS* **20**, 448–454.

- Daviet, L. & Schalk, M. (2010).** Biotechnology in plant essential oil production: progress and perspective in metabolic engineering of the terpene pathway. *Flavour Fragr J* **25**, 123–127.
- Delvigne, F., Baert, J., Gofflot, S., Lejeune, A., Telek, S., Johanson, T. & Lantz, A. E. (2014).** Dynamic single-cell analysis of *Saccharomyces cerevisiae* under process perturbation: Comparison of different methods for monitoring the intensity of population heterogeneity. *J Chem Technol Biotechnol* **90**, 314–323.
- Dhar, N. & McKinney, J. D. (2007).** Microbial phenotypic heterogeneity and antibiotic tolerance. *Curr Opin Microbiol* **10**, 30–38.
- Dubnau, D. (1991).** The regulation of genetic competence in *Bacillus subtilis*. *Mol Microbiol* **5**, 11–18.
- Eldar, A. & Elowitz, M. B. (2010).** Functional roles for noise in genetic circuits. *Nature* **467**, 167–173. Nature Publishing Group.
- Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. (2002).** Stochastic gene expression in a single cell. *Science* **297**, 1183–1186.
- Enfors, S. O., Jahic, M., Rozkov, a, Xu, B., Hecker, M., Jürgen, B., Krüger, E., Schweder, T., Hamer, G. & other authors. (2001).** Physiological responses to mixing in large scale bioreactors. *J Biotechnol* **85**, 175–85.
- Facklam, R. & Elliot, J. a. (1995).** Identification, classification, and clinical relevance of catalase negative, gram positive Cocci, excluding the Streptococci and Enterococci. *ClinMicrobiolRev* **8**, 479–495.
- Fernández-Castané, A., Vine, C. E., Caminal, G. & López-Santín, J. (2012).** Evidencing the role of lactose permease in IPTG uptake by *Escherichia coli* in fed-batch high cell density cultures. *J Biotechnol* **157**, 391–398. Elsevier B.V.
- Fey, P. D., Endres, J. L., Yajjala, V. K., Widhelm, T. J., Boissy, R. J., Bose, J. L. & Bayles, K. W. (2013).** A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. *MBio* **4**, 1–8.
- Figueiredo, A. M., Ha, E., Kreiswirth, B. N., de Lencastre, H., Noel, G. J., Senterfit, L. & Tomasz, A. (1991).** In vivo stability of heterogeneous expression classes in clinical isolates of methicillin-resistant staphylococci. *J Infect Dis* **164**, 883–887.
- García-Cayuela, T., de Cadiñanos, L. P. G., Mohedano, M. L., de Palencia, P. F., Boden, D., Wells, J., Peláez, C., López, P. & Requena, T. (2012).** Fluorescent protein vectors for promoter analysis in lactic acid bacteria and *Escherichia coli*. *Appl Microbiol Biotechnol* **96**, 171–181.
- Gasson, M. J. (1983).** Plasmid complements of *Streptococcus lactis* NCDO 712 and other Lactic Streptococci after protoplast-induced curing. *J Bacteriol* **154**, 1–9.



- Geoffroy, M.-C., Guyard, C., Quatannens, B., Pavan, S., Lange, M. & Mercenier, A. (2000).** Use of green fluorescent protein to tag lactic acid bacterium strains under development as live vaccine vectors. *Appl Environ Microbiol* **66**, 383–391.
- Georgopoulos, C. & Welch, W. J. (1993).** Role of the major heat shock proteins as molecular chaperones. *Annu Rev Cell Biol* **9**, 601–634.
- Givskov, M., Eberl, L., Møller, S., Poulsen, L. K. & Molin, S. (1994).** Responses to nutrient starvation in *Pseudomonas putida* KT2442: Analysis of general cross-protection, cell shape, and macromolecular content. *J Bacteriol* **176**, 7–14.
- Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. (2005).** Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–36.
- Gottesman, S. (1996).** Proteases and their targets in *Escherichia coli*. *Annu Rev Genet* **30**, 465–506.
- Gottesman, S., Wickner, S. & Maurizi, M. R. (1997).** Protein quality control: Triage by chaperones and proteases. *Genes Dev* **11**, 815–823.
- Goulian, M. & van der Woude, M. (2006).** A simple system for converting *lacZ* to *gfp* reporter fusions in diverse bacteria. *Gene* **372**, 219–226.
- Gourse, R. L., Gaal, T., Bartlett, M. S. J., Appleman, A. & Ross, W. (1996).** rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol* **50**, 645–77.
- Han, S., Delvigne, F., Brognaux, A., Charbon, G. E. & Sørensen, S. J. (2013).** Design of growth-dependent biosensors based on destabilized GFP for the detection of physiological behavior of *Escherichia coli* in heterogeneous bioreactors. *Biotechnol Prog* **29**, 553–563.
- Hansen, F. G. & Atlung, T. (2011).** YGFP: A spectral variant of GFP protein. *Biotechniques* **50**, 411–412.
- Hanson, G. T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R. a., Tsien, R. Y. & Remington, S. J. (2004).** Investigating Mitochondrial Redox Potential with Redox-sensitive Green Fluorescent Protein Indicators. *J Biol Chem* **279**, 13044–13053.
- Hartman, B. J. & Tomasz, A. (1986).** Expression of methicillin resistance in Heterogeneous Strains of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **29**, 85–92.
- Hartwell, L. H. & Unger, M. W. (1977).** Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. *J Cell Biol* **75**, 422–435.
- Hebisch, E., Knebel, J., Landsberg, J., Frey, E. & Leisner, M. (2013).** High Variation of Fluorescence Protein Maturation Times in Closely Related *Escherichia coli* Strains. *PLoS One* **8**.
- Hecker, M., Schumann, W. & Völker, U. (1996).** Heat-shock and general stress response in *Bacillus subtilis*. *Mol Microbiol* **19**, 417–28.

- Hedal, M., Norland, S., Bratback, G. & Riemann, B. (1994).** Determination of bacterial cell number and cell volume by means of flow cytometry, transmission electron microscopy, and epifluorescence microscopy. *J Microbiol Methods* **20**, 255–263.
- Heim, R. & Tsien, R. Y. (1996).** Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol* **6**, 178–82.
- Heim, R., Prasher, D. C. & Tsien, R. Y. (1994).** Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* **91**, 12501–4.
- Hewitt, C. J., Onyeaka, H., Lewis, G., Taylor, I. W. & Nienow, A. W. (2006).** A comparison of high cell density fed-batch fermentations involving both induced and non-induced recombinant *Escherichia coli* under well-mixed small-scale and simulated poorly mixed large-scale conditions. *Biotechnol Bioeng* **96**, 495–505.
- Hoch, J. A. (1993).** Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* **47**, 441–465.
- Holo, H. & Nes, I. F. (1989).** High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. *Appl Environ Microbiol* **55**, 3119–3123.
- Holzappel, W. H. & Wood, B. J. B. (2014).** *Lactic Acid Bacteria: Biodiversity and Taxonomy*. Wiley-Blackwell.
- Hutkins, R. W. & Nannen, N. L. (1993).** pH Homeostasis in Lactic Acid Bacteria. *J Dairy Sci* **76**, 2354–2365.
- Iizuka, R., Yamagishi-Shirasaki, M. & Funatsu, T. (2011).** Kinetic study of de novo chromophore maturation of fluorescent proteins. *Anal Biochem* **414**, 173–178. Elsevier Inc.
- Ingraham, J. L., Maaløe, O. & Neidhardt, F. C. (1983).** *Growth of the Bacterial Cell*. Sunderland: Sinauer Associates Inc.
- Ishihama, A. (1997).** Adaptation of gene expression in stationary phase bacteria. *Curr Opin Genet Dev* **7**, 582–588.
- Ishihama, A. (1999).** Modulation of the nucleoid, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival. *Genes to Cells* **4**, 135–143.
- Jagadish, M. N. & Carter, B. L. . (1977).** Genetic control of cell division in yeast cultured at different growth rates. *Nature* **269**, 145–147.
- Jagadish, M. N., Lorincz, A. & Carter, B. L. . (1977).** Cell size and cell division in yeast cultured at different growth rates. *FEMS Microbiol Lett* **2**, 235–237.
- Jenkins, D., Schultz, J. & Matin, A. (1990).** Starvation-induced cross protection against osmotic challenge in *Escherichia coli*. *J Bacteriol* **172**, 2779–2781.

- Jensen, P. R., Westerhoff, H. V & Michelsen, O. (1993).** The use of *lac*-type promoters in control analysis. *Eur J Biochem* **211**, 181–91.
- Jensen, P. R. & Hammer, K. (1993).** Minimal requirements for exponential growth of *Lactococcus lactis*. *Appl Environ Microbiol* **59**, 4363–4366.
- John, R. P., Nampoothiri, K. M. & Pandey, A. (2007).** Fermentative production of lactic acid from biomass: An overview on process developments and future perspectives. *Appl Microbiol Biotechnol* **74**, 524–534.
- Johnson, B. F. (1965).** Morphometric analysis of yeast cells. *Exp Cell Res* **39**, 577–583.
- De Jong, I. G., Veening, J.-W. & Kuipers, O. P. (2012).** Single cell analysis of gene expression patterns during carbon starvation in *Bacillus subtilis* reveals large phenotypic variation. *Environ Microbiol* **14**, 3110–3121.
- Julià, O., Comas, J. & Vives-Rego, J. (2000).** Second-order functions are the simplest correlations between flow cytometric light scatter and bacterial diameter. *J Microbiol Methods* **40**, 57–61.
- Julià, O. & Vives-Rego, J. (2005).** Skew-Laplace distribution in Gram-negative bacterial axenic cultures: New insights into intrinsic cellular heterogeneity. *Microbiology* **151**, 749–755.
- Julià, O., Vidal-Mas, J., Panikov, N. S. & Vives-Rego, J. (2010).** Skew-laplace and cell-size distribution in microbial axenic cultures: statistical assessment and biological interpretation. *Int J Microbiol* **2010**, 191585.
- Kacmar, J., Zamamiri, A., Carlson, R., Abu-Absi, N. R. & Srienc, F. (2004).** Single-cell variability in growing *Saccharomyces cerevisiae* cell populations measured with automated flow cytometry. *J Biotechnol* **109**, 239–54.
- Kaempfer, R. O. & Magasanik, B. (1967).** Mechanism of beta-galactosidase induction in *Escherichia coli*. *J Mol Biol* **27**, 475–494.
- Kearns, D. B. & Losick, R. (2005).** Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev* 3083–3094.
- Keiler, K. C. & Sauer, R. T. (1996).** Sequence determinants of C-terminal substrate recognition by the Tsp protease. *J Biol Chem* **271**, 2589–93.
- Keren, L., Dijk, D. Van, Weingarten-gabbay, S. & Davidi, D. (2015).** Noise in gene expression is coupled to growth rate. *Genome Res*.
- Kesel, S., Mader, A., Höfler, C., Mascher, T. & Leisner, M. (2013).** Immediate and Heterogeneous Response of the LiaFSR Two-Component System of *Bacillus subtilis* to the Peptide Antibiotic Bacitracin. *PLoS One* **8**, e53457.

- Kim, C., Mwangi, M., Chung, M., Milheirco, C., De Lencastre, H. & Tomasz, A. (2013).** The mechanism of heterogeneous beta-lactam resistance in MRSA: Key role of the stringent stress response. *PLoS One* **8**, 1–10.
- Kjeldgaard, N. O., Maaloe, O. & Schaechter, M. (1958).** The transition between different physiological states during balanced growth of *Salmonella typhimurium*. *J Gen Microbiol* **19**, 607–616.
- Kohanski, M. a., Dwyer, D. J., Hayete, B., Lawrence, C. a. & Collins, J. J. (2007).** A Common Mechanism of Cellular Death Induced by Bactericidal Antibiotics. *Cell* **130**, 797–810.
- Kolter, R., Siegele, D. a & Tormo, A. (1993).** The stationary phase of the bacterial life cycle. *Annu Rev Microbiol* **47**, 855–874.
- Le, T. T., Harlepp, S., Guet, C. C., Dittmar, K., Emonet, T., Pan, T. & Cluzel, P. (2005).** Real-time RNA profiling within a single bacterium. *Proc Natl Acad Sci U S A* **102**, 9160–9164.
- Levsky, J. M., Shenoy, S. M. & Pezo, R. C. (2002).** Single-Cell Gene Expression Profiling. *Science (80- )* **297**, 836–840.
- Lim, D. & Strynadka, N. C. J. (2002).** Structural basis for the beta lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus*. *Nat Struct Biol* **9**, 870–876.
- Lopez de Felipe, F., Kleerebezem, M., de Vos, W. M. & Hugenholtz, J. (1998).** Cofactor Engineering : a Novel Approach to Metabolic Engineering in *Lactococcus lactis* by Controlled Expression of NADH Oxidase. *J Bacteriol* **180**, 3804–3808.
- Lopez de Felipe, F., Starrenburg, M. J. . & Hugenholtz, J. (1997).** The role of NADH-oxidation in acetoin and diacetyl production from glucose in *Lactococcus lactis* subsp. *lactis* MG1363. *FEMS Microbiol Lett* **156**, 15–19.
- Madar, D., Dekel, E., Bren, A., Zimmer, A., Porat, Z. & Alon, U. (2013).** Promoter activity dynamics in the lag phase of *Escherichia coli*. *BMC Syst Biol* **7**, 136. BMC Systems Biology.
- Martínez, J. L., Liu, L., Petranovic, D. & Nielsen, J. (2012).** Pharmaceutical protein production by yeast: Towards production of human blood proteins by microbial fermentation. *Curr Opin Biotechnol* **23**, 965–971.
- Mcdougal, L. K., Steward, C. D., Killgore, G. E., Chaitram, J. M., Mcallister, S. K. & Tenover, F. C. (2003).** Pulsed-Field Gel Electrophoresis Typing of Oxacillin-Resistant. *J Clin Microbiol* **41**, 5113–5120.
- Mercade, M., Lindley, N. D. & Loubière, P. (2000).** Metabolism of *Lactococcus lactis* subsp. *cremoris* MG1363 in acid stress conditions. *Int J Food Microbiol* **55**, 161–165.
- Million-Weaver, S., Alexander, D. L., Allan, J. M. & Camps, M. (2012).** Methods for quantifying plasmid copy number to investigate plasmid dosage effects associated with directed protein evolution. *Methods Mol Biol* **834**, 33–48.

- Ming Bo Huang, Gay, T. E., Baker, C. N., Banerjee, S. N. & Tenover, F. C. (1993).** Two percent sodium chloride is required for susceptibility testing of staphylococci with oxacillin when using agar-based dilution methods. *J Clin Microbiol* **31**, 2683–2688.
- Miyawaki, a, Griesbeck, O., Heim, R. & Tsien, R. Y. (1999).** Dynamic and quantitative Ca<sup>2+</sup> measurements using improved cameleons. *Proc Natl Acad Sci U S A* **96**, 2135–2140.
- Miyoshi, A., Poquet, I., Azevedo, V., Commissaire, J., Domakova, E., Loir, Y. Le, Oliveira, S. C., Gruss, A. & Langella, P. (2002).** Controlled Production of Stable Heterologous Proteins in *Lactococcus lactis*. *Appl Environ Microbiol* **68**, 3141–3146.
- Mwangi, M. M., Kim, C., Chung, M., Tsai, J., Vijayadamodar, G., Benitez, M., Jarvie, T. P., Du, L. & Tomasz, A. (2013).** Whole-genome sequencing reveals a link between  $\beta$ -lactam resistance and synthetases of the alarmone (p)ppGpp in *Staphylococcus aureus*. *Microb Drug Resist* **19**, 153–9.
- Nannen, N. L. & Hutkins, R. W. (1991).** Intracellular pH effects in lactic acid bacteria. *J Dairy Sci* **74**, 741–746.
- Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. (1990).** *Physiology of the bacterial cell: a molecular approach*. Sunderland, Massachusetts: Sinauer Associates Inc.
- Neidhardt, F. C. & Magasanik, B. (1960).** Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim Biophys Acta* **42**, 99–116.
- Niu, Y., Kong, J. & Xu, Y. (2008).** A novel GFP-fused eukaryotic membrane protein expression system in *Lactococcus lactis* and its application to overexpression of an elongase. *Curr Microbiol* **57**, 423–428.
- Novick, A. & Weiner, M. (1957).** Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci* **43**, 553–566.
- Nyström, T. (2003).** The free-radical hypothesis of aging goes prokaryotic. *Cell Mol Life Sci* **60**, 1333–1341.
- Oddone, G. M., Lan, C. Q., Rawsthorne, H., Mills, D. A. & Block, D. E. (2007).** Optimization of fed-batch production of the model recombinant protein GFP in *Lactococcus lactis*. *Biotechnology* **96**, 1127–1138.
- Oddone, G. M., Mills, D. a & Block, D. E. (2009).** Dual inducible expression of recombinant GFP and targeted antisense RNA in *Lactococcus lactis*. *Plasmid* **62**, 108–118. Elsevier Inc.
- Otto, M. (2010).** Basis of virulence in community-associated methicillin-resistant *Staphylococcus aureus*. *Annu Rev Microbiol* **64**, 143–162.
- Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. (2002).** Regulation of noise in the expression of a single gene. *Nat Genet* **31**, 69–73.
- Paterson, G. K., Harrison, E. M. & Holmes, M. a. (2014).** The emergence of mecC methicillin-resistant *Staphylococcus aureus*. *Trends Microbiol* **22**, 42–47. Elsevier Ltd.

- Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. & Piston, D. W. (1997).** Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* **73**, 2782–90. Elsevier.
- Paulander, W., Wang, Y., Folkesson, A., Charbon, G., Løbner-Olesen, A. & Ingmer, H. (2014).** Bactericidal antibiotics increase hydroxyphenyl fluorescein signal by altering cell morphology. *PLoS One* **9**, e92231.
- Pedersen, A. E., Thyregod, C., Kilstrup, M. & Martinussen, J. (2015).** Heterogeneity at a single cell level in homogenous cultures. *in preparation*.
- Piersma, S., Denham, E. L., Drulhe, S., Tonk, R. H. J., Schwikowski, B. & van Dijl, J. M. (2013).** TLM-Quant: An Open-Source Pipeline for Visualization and Quantification of Gene Expression Heterogeneity in Growing Microbial Cells. *PLoS One* **8**, 1–6.
- Pierucci, O. (1978).** Dimensions of *Escherichia coli* at various growth rates : model for envelope growth. *J Bacteriol* **135**, 559–574.
- Pinto, J. P. C., Zeyniyev, A., Karsens, H., Trip, H., Lolkema, J. S., Kuipers, O. P. & Kok, J. (2011).** pSEUDO, a genetic integration standard for *Lactococcus lactis*. *Appl Environ Microbiol* **77**, 6687–6690.
- Prasher, D. C. (1995).** Using GFP to see the light. *Trends Genet* **11**, 320–3.
- Prüb, B. M., Nelms, J. M., Park, C. & Wolfe, A. J. (1994).** Mutations in NADH:Ubiquinone oxidoreductase of *Escherichia coli* affect growth on mixed amino acids. *J Bacteriol* **176**, 2143–2150.
- Rawsthorne, H., Turner, K. N. & Mills, D. a. (2006).** Multicopy integration of heterologous genes, using the lactococcal group II intron targeted to bacterial insertion sequences. *Appl Environ Microbiol* **72**, 6088–6093.
- Reed, S. I. (1980).** The selection of *S. cerevisiae* mutants defective in the start event of cell division. *Genetics* **95**, 561–577.
- Regenberg, B., Grotkjaer, T., Winther, O., Fausbøll, A., Akesson, M., Bro, C., Hansen, L. K., Brunak, S. & Nielsen, J. (2006).** Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. *Genome Biol* **7**, R107.
- Rizzo, M. a, Springer, G. H., Granada, B. & Piston, D. W. (2004).** An improved cyan fluorescent protein variant useful for FRET. *Nat Biotechnol* **22**, 445–449.
- Robertson, B. R., Button, D. K. & Koch, a. L. (1998).** Determination of the biomasses of small bacteria at low concentrations in a mixture of species with forward light scatter measurements by flow cytometry. *Appl Environ Microbiol* **64**, 3900–3909.
- Rodríguez, J. M., Martínez, M. I., Horn, N. & Dodd, H. M. (2003).** Heterologous production of bacteriocins by lactic acid bacteria. *Int J Food Microbiol* **80**, 101–16.

- Rolfe, M. D., Rice, C. J., Lucchini, S., Pin, C., Thompson, A., Cameron, A. D. S., Alston, M., Stringer, M. F., Betts, R. P. & other authors. (2012). Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *J Bacteriol* **194**, 686–701.
- Roszak, D. B. & Colwell, R. R. (1987). Survival strategies of bacteria in the natural environment. *Microbiol Rev* **51**, 365–379.
- Saint-Ruf, C., Pesut, J., Sopta, M. & Matic, I. (2008). Causes and consequences of DNA repair activity modulation during stationary phase in *Escherichia coli*. *Crit Rev Biochem Mol Biol* **42**, 259–270.
- Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning A Laboratory Manual*.
- Sargent, M. G. (1975). Control of Cell Length in *Bacillus subtilis*. *J Bacteriol* **123**, 7–19.
- Sastalla, I., Chim, K., Cheung, G. Y. C., Pomerantsev, A. P. & Leppla, S. H. (2009). Codon-optimized fluorescent proteins designed for expression in low-GC gram-positive bacteria. *Appl Environ Microbiol* **75**, 2099–2110.
- Schaechter, M., Maaloe, O. & Kjeldgaard, N. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *Microbiology* **19**, 592–606.
- Schaechter, M., Williamson, J. P., Hood, J. R. & Koch, a L. (1962). Growth, cell and nuclear divisions in some bacteria. *J Gen Microbiol* **29**, 421–434.
- Scholz, O., Thiel, A., Hillen, W. & Niederweis, M. (2000). Quantitative analysis of gene expression with an improved green fluorescent protein. *Eur J Biochem* **267**, 1565–1570.
- Schumann, W. (2003). The *Bacillus subtilis* heat shock stimulon. *Cell Stress Chaperones* **8**, 207217.
- Scott, K. P., Mercer, D. K., Richardson, a J., Melville, C. M., Glover, L. a & Flint, H. J. (2000). Chromosomal integration of the green fluorescent protein gene in lactic acid bacteria and the survival of marked strains in human gut simulations. *FEMS Microbiol Lett* **182**, 23–7.
- Sezonov, G., Joseleau-Petit, D. & D'Ari, R. (2007). *Escherichia coli* physiology in Luria-Bertani broth. *J Bacteriol* **189**, 8746–8749.
- Sharpe, M. E., Hauser, P. M., Sharpe, R. G. & Errington, J. (1998). *Bacillus subtilis* cell cycle as studied by fluorescence microscopy: Constancy of cell length at initiation of DNA replication and evidence for active nucleoid partitioning. *J Bacteriol* **180**, 547–555.
- Siemering, K. R., Golbik, R., Sever, R. & Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr Biol* **6**, 1653–1663.
- Smith, C. (2002). Stability of green fluorescent protein using luminescence spectroscopy: is GFP applicable to field analysis of contaminants? *Environ Pollut* **120**, 517–520.

- Solomon, J. M. & Grossman, A. D. (1996).** Who's competent and when: Regulation of natural genetic competence in bacteria. *Trends Genet* **12**, 150–155.
- Strovas, T. J. & Lidstrom, M. E. (2009).** Population heterogeneity in *Methylobacterium extorquens* AM1. *Microbiology* **155**, 2040–8.
- Swain, P. S., Elowitz, M. B. & Siggia, E. D. (2002).** Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* **99**, 12795–800.
- Tachon, S., Chambellon, E. & Yvon, M. (2011).** Identification of a conserved sequence in flavoproteins essential for the correct conformation and activity of the NADH oxidase NoxE of *Lactococcus lactis*. *J Bacteriol* **193**, 3000–3008.
- Thyregod, C., Pedersen, A. E., Kilstrup, M. & Martinussen, J. (2015).** Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry. *in preparation*.
- Tomasz, A. & Waks, S. (1975).** Mechanism of action of penicillin: Triggering of the pneumococcal autolytic enzyme by inhibitors of cell wall synthesis. *Proc Natl Acad Sci* **72**, 4162–4166.
- Tomasz, A., Nachman, S. & Leaf, H. (1991).** Stable Classes of Phenotypic Expression in Methicillin-Resistant Clinical Isolates of Staphylococci. *Antimicrob Agents Chemother* **35**, 124–129.
- Tsien, R. Y. (1998).** The green fluorescent protein. *Annu Rev Biochem* **67**, 509–544.
- Tyson, C. B., Lord, P. G. & Wheals, a. E. (1979).** Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. *J Bacteriol* **138**, 92–98.
- Umakoshi, H., Tanabe, T., Suga, K., Bui, H. T., Shimanouchi, T. & Kuboi, R. (2011).** Oxidative stress can affect the gene silencing effect of DOTAP liposome in an *in vitro* translation system. *Int J Biol Sci* **7**, 253–260.
- Veening, J.-W., Smits, W. K. & Kuipers, O. P. (2008).** Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**, 193–210.
- Verduyn, C., Postma, E., Scheffers, A. & Van Dijken, J. P. (1992).** Effect of Benzoic Acid on Metabolic Fluxes in Yeasts : A Continuous-Culture Study on the Regulation of Respiration and Alcoholic Fermentation. *Yeast* **8**, 501–517.
- Vessoni Penna, T. C., Ishii, M., Cholewa, O. & De Souza, L. C. (2004).** Thermal characteristics of recombinant green fluorescent protein (GFPuv) extracted from *Escherichia coli*. *Lett Appl Microbiol* **38**, 135–139.
- Vives-Rego, J., Resina, O., Comas, J., Loren, G. & Julià, O. (2003).** Statistical analysis and biological interpretation of the flow cytometric heterogeneity observed in bacterial axenic cultures. *J Microbiol Methods* **53**, 43–50.



- Wada, a, Yamazaki, Y., Fujita, N. & Ishihama, A. (1990).** Structure and probable genetic location of a 'ribosome modulation factor' associated with 100S ribosomes in stationary-phase *Escherichia coli* cells. *Proc Natl Acad Sci U S A* **87**, 2657–2661.
- Watson, S. P., Clements, M. O. & Foster, S. J. (1998).** Characterization of the starvation-survival response of *Staphylococcus aureus*. *J Bacteriol* **180**, 1750–1758.
- Wehr, C. T. & Parks, L. W. (1969).** Macromolecular synthesis in *Saccharomyces cerevisiae* in different growth media. *J Bacteriol* **98**, 458–466.
- Williams, T. C., Espinosa, M. I., Nielsen, L. K. & Vickers, C. E. (2015).** Dynamic regulation of gene expression using sucrose responsive promoters and RNA interference in *Saccharomyces cerevisiae*. *Microb Cell Fact* **14**, 1–10.
- Wyszyńska, A., Kobierecka, P., Bardowski, J. & Jagusztyn-Krynicka, E. K. (2015).** Lactic acid bacteria—20 years exploring their potential as live vectors for mucosal vaccination. *Appl Microbiol Biotechnol* **99**, 2967–2977.
- Yilmaz, S. & Singh, A. K. (2012).** Single cell genome sequencing. *Curr Opin Biotechnol* **23**, 437–443.
- Zhang, K., Martiny, A. C., Reppas, N. B., Barry, K. W., Malek, J., Chisholm, S. W. & Church, G. M. (2006a).** Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol* **24**, 680–686.
- Zhang, L., Patel, H. N., Lappe, J. W. & Wachter, R. M. (2006b).** Reaction progress of chromophore biogenesis in green fluorescent protein. *J Am Chem Soc* **128**, 4766–4772.

## 11 Appendix

### 11.1 Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry

Camilla Thyregod<sup>2</sup>, Anne Egholm Pedersen<sup>1</sup>, Mogens Kilstrup<sup>1</sup>, and Jan Martinussen<sup>1\*</sup>

<sup>1</sup>Metabolic Signaling and Regulation Group, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>2</sup>: Institut for Matematik og Computer Science, DTU compute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>\*\*</sup>Corresponding author. Mailing address:

Metabolic Signaling and Regulation Group, Department of Systems Biology, Building 301,  
Technical University of Denmark, 2800 Kgs. Lyngby, Denmark  
Phone: +45 45252498, E-mail: jma@bio.dtu.dk

Keywords: Population heterogeneity, box-plot, qq-plot, mixture of subpopulations, descriptive statistics, GFP, flow cytometry, lognormal, log-skew laplace.

Keywords: Population heterogeneity, GFP, flow cytometry, balanced growing cultures

### 11.1.1 Abstract

Population heterogeneity determined on a single cell level can be elucidated by flow cytometry analysis. Dot plot, density plots or similar offers the ability to make a subjective assessment of the degree of heterogeneity within a sample and between samples. Statistical methods can be used to obtain a more quantitative and objective determination of the degree of heterogeneity of flow cytometry data. Modelling of distributions to the flow cytometry data can show the presence of subpopulations and identify the similarity of the heterogeneity/distribution of different samples. Comparing heterogeneity of large samples following different types of statistical distributions may not be straight forward and thus descriptive statistics in combination with other simple statistical methods presented can elucidate the degree of heterogeneity within and between samples of flow cytometry data, without consideration about sample distribution. Descriptive statistics is in this paper used to show that the degree of heterogeneity in single cell growth rate is lowest in a balanced growing *B. subtilis* culture, compared to the degree observed in a *B. subtilis* culture inoculated with a stationary phase pre-culture.

### 11.1.2 Introduction

Flow cytometry is a powerful single cell analysis tool, which makes it possible to analyze many cells per second. Information about cell size and fluorescence, whether being fluorescence staining or fluorescence produced by the cell by expression of a gene encoding a fluorescent protein, can be collected for each single cell in a sample. Therefore, flow cytometry can be used to report heterogeneity in a sample. When describing heterogeneity in relation to single cell analysis by flow cytometry it is relevant to consider different aspects of heterogeneity like variation, location as well as distribution, including the existence of subpopulations.

In the literature population heterogeneity is often discussed based on dot plots (or similar) leaving conclusions on heterogeneity to be made based on rather subjective assessments. In order to introduce a more objective input to the discussion of heterogeneity a number of articles based in statistical theory have been published. The focus is on finding a suitable family of distributions that generally fit to flow cytometry data under various experimental conditions. Various probability distributions have been suggested and investigated in relation to particle size data generally and (based on these investigations) more specifically in relation to flow cytometry data. In previous work various probability distributions have been suggested and investigated in relation to particle size data generally and (based on these investigations) more specifically in relation to flow cytometry data. The log-normal distribution is a standard distribution in size (e.g. particle size) statistics (Barndorff-Nielsen, 1977). When size follows a log-normal distribution it is equivalent to the log-transformed size following a normal distribution. However, often the logarithm of the size does not follow a normal distribution. Barndorff-Nielsen therefore introduces and investigates the class of hyperbolic probability distributions. When the logarithm of the size follows a hyperbolic distribution, it is referred to as a log-hyperbolic distribution (because of the initial logarithm transformation of size). The log-normal distribution is a limiting form of the log-hyperbolic function, and moreover the log-hyperbolic function can be generated as a

mixture of log-normal distributions. In relation to flow cytometry this is relevant as the distribution seen in a sample could be the overall distribution of a population actually consisting of a mixture of two or more subpopulations. Generally, this means that if each subpopulation follows a log-normal distribution, the overall distribution of the mixture of subpopulations can often be described by a log-hyperbolic distribution. The theory and applications of hyperbolic distributions are summarized and further investigated in (Bagnold & Barndorff-Nielsen, 1980). Fieller et al. argues in favor of the log-skew-Laplace distribution: the log-skew-Laplace distribution is a limiting distribution of the family of log-hyperbolic distributions (Fieller *et al.*, 1992). The log-skew-Laplace distribution is simpler to work with compared to the log-hyperbolic distributions and yet the fit is not appreciably different from that of the log-hyperbolic model. The log-skew-Laplace model provides a better fit to particle size data from samples of sand compared to the log-normal model, and further the log-skew-Laplace model can (similarly to the log-hyperbolic model) be used in the situation where the overall distribution is the result of a mixture of the distributions for subpopulations. Kotz et al. elaborates the work with the Laplace distribution (Kotz *et al.*, 2001). They described e.g. that a classical symmetric Laplace random variable may be viewed as a normal random variable with mean zero and a stochastic variance. Analogously a skewed-Laplace random variable has a similar interpretation, where the mean of the normal distribution is now also stochastic. I.e. in relation to flow cytometry data this informally means that if we have a mixture of subpopulations where the distribution of each subpopulation is log-normal (or normal), and where the mean and variances are both random variables (following a specified probability distribution themselves) the overall distribution for the mixture of subpopulations can be described by a log-skew-Laplace (skew-Laplace) distribution.

Julià and Vives-Rego applies the general ideas of modelling particle size by the use of the (log)-skew-Laplace distribution specifically to flow cytometry data (Julià & Vives-Rego, 2005). They found that SSC values fits with both distributions. As the skew-Laplace distribution can be thought of as a mixture of normal distributions it is suggested that the results for the SSC values shows that analyzed Gram-negative cultures may consist of a number of statistically independent subpopulations.

Julia et al. investigated in 2010 the fitting of the skew-Laplace to FSC, SSC and Multisizer cell size, for different microorganisms (Julià *et al.*, 2010). They found that most of the cases, where the skew-Laplace fit is satisfactory, correspond to the exponential phase and for microorganisms with low morphological variability.

A possible explanation of why the skew-Laplace fit particularly seemed to be satisfactory for the exponential growth phase can be found in the work by Prats et al. (Prats *et al.*, 2008, 2010). In the article from 2008 Prats et al. modelled stages in the bacterial lag phase by handling the evolution of the total biomass and the total number of bacteria separately. Mathematical simulations based on inoculated bacteria sampled from a previous simulation of a culture in the stationary phase showed that the initial mean mass of the population was small, and the distribution of the biomass differed from the corresponding distribution found for the exponential phase. During the lag phase there is an increase in mean biomass and the shape of the distribution of biomass changes showing a forward shift (i.e. an increasing tail of the distribution to the right).

In the exponential phase the biomass distribution is stable. Prats et al. also show by mathematical simulations that when nutrient run short and the culture enters stationary phase, the shape of the distribution of the biomass shifts backwards and the mean mass decreases. The mathematical based results from 2008 are validated experimentally for FSC (transformed to cell diameter) in an *E. coli* batch culture (Prats et al., 2010).

In conclusion the literature overview showed that several probability distributions have been suggested for describing the heterogeneity of cellular properties in an isogenic culture. Lately particularly the (log)-skew-Laplace distribution has received focus. This distribution is closely related to a mixture of (log)-normal distributions indicating that in some situations the overall distribution of a cellular property is the result of a mixture of the distributions describing individual subpopulations regarding that specific cellular property. The (log)-skew-Laplace distribution seems to be particularly relevant in the exponential growth phase for a population. However, factors as the type of cell, the growth phase etc. may change the distribution of cell size (Julià et al., 2010).

As most non-statisticians working with flow cytometry may be familiar with distributions like the normal and the log-normal distributions but not with distributions like the (log)-skew-Laplace distribution, and as further the relevant distribution for describing sample heterogeneity for a cell culture may depend on the growth phase complicating things further, there is a need for input on how to assess heterogeneity using relatively simple statistical methods: even though simple, assessing heterogeneity using these methods will lead to more objective conclusions compared to what is often done in the literature at the present.

In this paper we investigate by flow cytometry the heterogeneity of an isogenic culture (in biological terms: homogeneous) of a *B. subtilis* GFP reporter strain. Expression of GFP in this reporter strain is placed under regulation of a ribosomal promoter and we show in adjacent paper that the GFP activity is a measure of growth rate (Pedersen et al. 2015). Hence, the distribution of the GFP activity of single cells is reflecting the distribution of growth rates of the individual cells. In the present work the statistical analyses of the experimental results follow the ideas and results outlined above, however from a more heuristic approach. The aim with the approach is to suggest guidelines for a non-statistician working with flow cytometry on how to analyze data. The statistical analyses are made using the statistical point-and-click software JMP®.

### 11.1.3 Materials and Methods

**Strains and plasmids** see (Pedersen et al., 2015)

**Media.** *B. subtilis* was grown in Spizizen salt medium supplemented with 1 mg of thiamine and 40 mg tryptophan per liter, as previously described (Saxild & Nygaard, 1987). Carbon sources were added to a final

concentration of 0.4 %. This medium will be referred to as chemically defined medium. The strains were cultivated in Erlenmeyer flasks at 37° C with vigorous shaking.

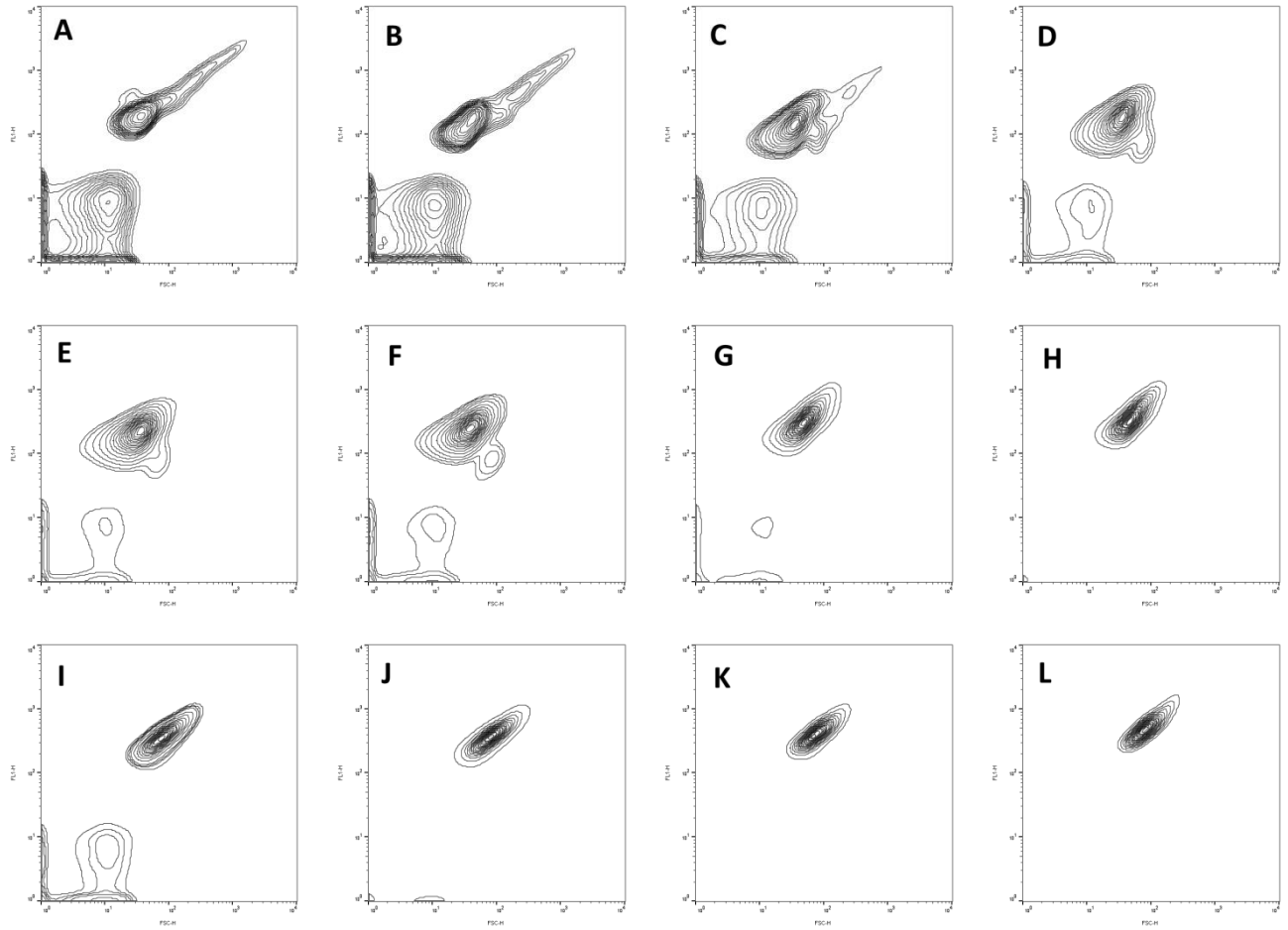
**Flow cytometry.** For flow cytometric analysis the cells were diluted in PBS solution: NaCl, 0.137 M; KCl, 3 mM; NaHPO<sub>4</sub>, 12 mM; and KH<sub>2</sub>PO<sub>4</sub>, 2 mM; pH 7.4, and measured directly on a BD FACSCalibur (Becton-Dickinson, NJ, USA) flow cytometer. Excitation wavelength for the laser used was 488 nm. Fluorescence emission levels were measured using a band pass filter at 530/30 nm (FITC). For each sample approximately 100,000 cells were analyzed. Contour plots for visualization of flow cytometry data were made in FlowJo. The contour plots used have 5 % probability contours, meaning that 5 % of the cells lie between each contour.

**Gating and setting forward scatter threshold before statistical analysis of flow cytometry data.** Data obtained from flow cytometry analysis was histogram gated based on fluorescence per cell, to isolate GFP containing cells from non-viable cells with permeable wall and cell debris. Flowing Software (v2.5.1, Turku Centre for Biotechnology, Turku, Finland) was used for gating. After gating, the fluorescence per forward scatter (specific cellular fluorescence) was calculated for each event. The relevance of this specific parameter is described in (Pedersen *et al.*, 2015). The statistical analysis of specific cellular fluorescence was made in the statistical software JMP® version 11.

#### 11.1.4 Results

##### 11.1.4.1 Population heterogeneity in chemically defined medium cultivations.

Flow cytometry analysis of the *B. subtilis* GFP reporter strain (AE099) in two different experimental setups is included to illustrate how statistical analysis of flow cytometry data can be used to investigate population heterogeneity in a culture of isogenic cells. The reporter strain was propagated in chemically defined medium at 37 °C with vigorously shaking. Samples were centrifuged and resuspended in ice cold PBS solution before analysis. The two experimental setups were attained by using two inoculums at different growth states; a stationary phase pre-culture and an exponentially growing pre-culture, both propagated in chemically defined medium. The culture inoculated with the exponentially growing pre-culture was defined as a balanced growing culture, in which the population heterogeneity was expected to be lowered to a minimum (Pedersen *et al.*, 2015). The flow cytometry analyses of the two resulting cultivations are shown in Figure 1 as contour plots. The culture inoculated with the stationary phase pre-culture was in the lag- and exponential growing phase for a longer time than the culture inoculated with the exponential growing pre-culture. Therefore more samples corresponding to the stationary phase pre-culture is included in Figure 1.



**Figure 1** Flow cytometry data obtained from analysis of *B. subtilis* GFP reporter strain AE099, presented as contour plots with 5 % probability contours. Horizontal axis: Forward scatter [FSC-H], log scale four decades. Vertical axis: Fluorescence [FL1-H], log scale four decades. A-H *B. subtilis* GFP reporter strain cultivated in chemically defined medium, inoculated with a stationary phase pre-culture. The samples were taken 12, 112, 217, 330, 397, 483, 632 and 762 minutes after inoculation. The growth curve can be seen above the box-plots in Figure 2A. I-L *B. subtilis* GFP reporter strain cultivated in chemically defined medium, inoculated with an exponential growing pre-culture, the four samples was taken 31, 98, 250 and 356 minutes after inoculation, the growth curve can be seen above the box-plots in Figure 2B.

The fluorescent cells were gated (see materials and methods) and fluorescence and forward scatter (proportional to cell size) were used to compute specific cellular fluorescence for each detected event. Specific cellular fluorescence is a measure of the ribosomal promoter regulated ( $P_{rm-gfp}$ ) expression per cell, which is proportional to the growth rate of the individual cells (Pedersen *et al.*, 2015)

#### 11.1.4.2 Defining heterogeneity

When working with flow cytometry data it would be practical if the heterogeneity of an individual sample could be expressed as a single heterogeneity index. Calculating the numerical value of such an index for a number of samples would be an easy way to compare the heterogeneity between samples with the aim of objectively identifying e.g. the most heterogeneous sample.

However, heterogeneity is a complex concept in relation to flow cytometry data and therefore so far a single index covering all aspects of heterogeneity related to flow cytometry data has (to our knowledge) not been developed.

Different aspects of heterogeneity in relation to flow cytometry data are variation, location (e.g. mean) as well as distribution, including the existence of subpopulations within the individual sample. Further, it is sometimes relevant to compare how each of these aspects varies across different samples (eventually from different experimental conditions) or develop over time for a specific culture. As an example of the complexity of assessing the heterogeneity between samples, consider the case where three samples have similar overall variation and location (and therefore according to these terms can be said to be similar), but differ with respect to distribution: one sample consists of one population, whereas the two other samples are mixtures of subpopulations.

As a single heterogeneity index converting all these aspects into one single number, does not exist, we will in the following assess these aspects individually as objectively as possible using simple statistical methods.

#### 11.1.4.3 Descriptive statistics and box-plots

Simple descriptive statistics like the mean, median, interquartile range (IQR), standard deviation (SD) and coefficient of variation (CV) are useful quantitative measures in relation to describing different aspects of heterogeneity.

The mean and the median are both measures of location or central tendency. In a symmetrical distribution like the (theoretical) normal distribution the mean and the median are identical. However, if the data can be described by a skewed distribution the mean and the median will differ. In this case the median is often preferred in relation to assessing the location of the sample.

IQR, SD and CV are different measures of variation (or dispersion). IQR is a range that covers the middle 50% of the data (third quartile minus first quartile). The median and the IQR are corresponding measures of location and central tendency as they are both expressed in terms of quartiles. The SD is known by most practitioners as it is widely used in relation to confidence intervals, measurement uncertainty etc. CV is a standardized measure of the standard deviation as it is the ratio of the SD to the mean.

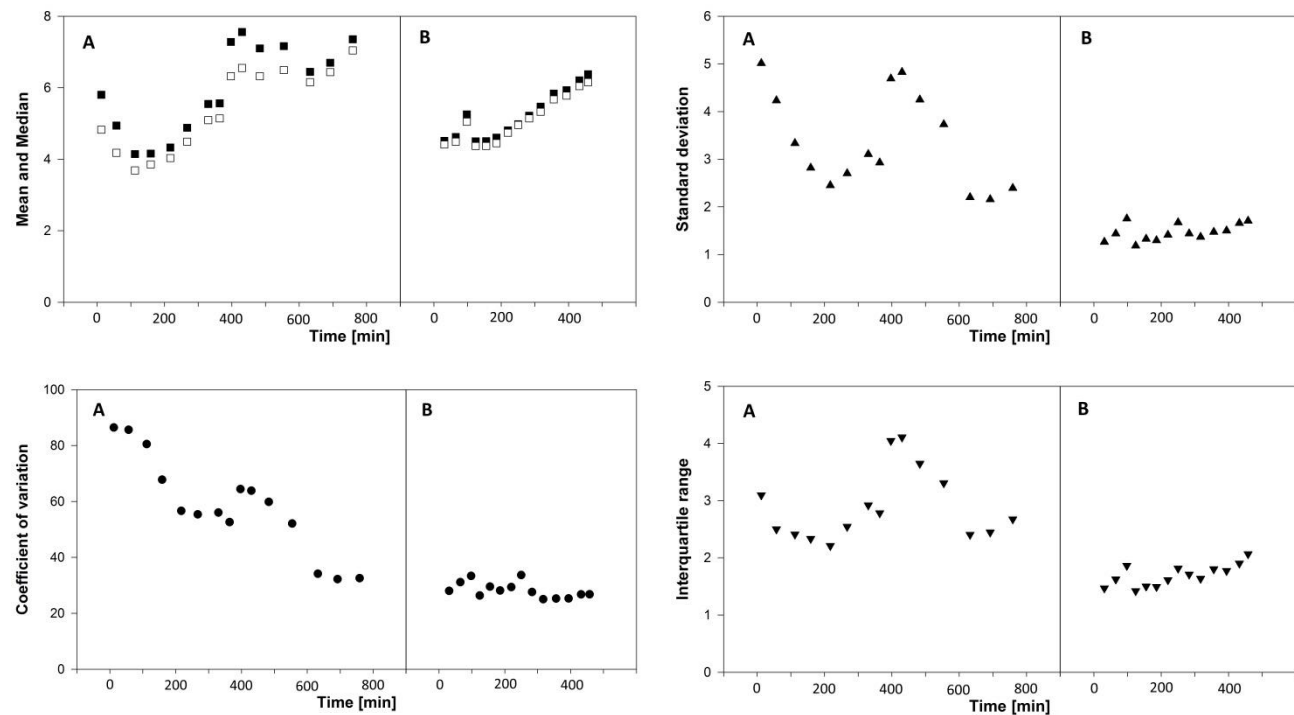
For a skewed distribution like the log-normal distribution it is an inherent property that the variance ( $SD^2$ ) is a function of the mean. When the mean increases it naturally follows that  $SD^2$  (and SD) increases. In relation to flow cytometry data this means that if the observations can be described by a log-normal distribution, the variation in a sample will naturally increase without any external factors causing the increase in the variation if



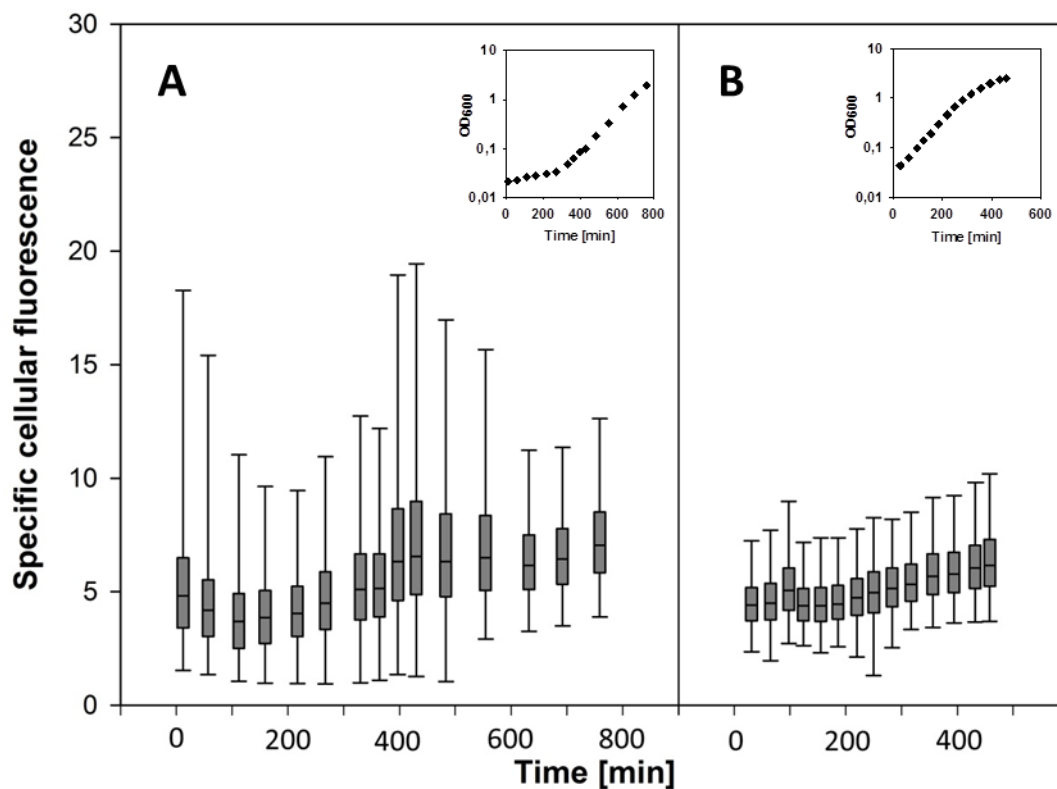
the mean is increased. Generally in relation to biological growth is often proportional to the number of cells. Therefore weight, height, length etc. of biological individuals are often described by a log-normal variable. In this situation it is relevant to consider the relative variation (CV).

Descriptive statistics can be calculated without making any assumptions regarding the distribution of the individual samples. Descriptive statistics are often presented in tables. In order to get an overview of the descriptive statistics, some of these can be summarized visually in so called box-plots. Box-plots provide an intuitive overview of: Location (mean and/or median), within sample and between sample variations as well as of the skewness of the distribution(s). Different versions of box-plots exist, however, for the box-plots presented in this article the bottom and top of the boxes are the first and third quartile of the data from the corresponding sample, the line dividing the box represents the median value: 50% of the observations in the sample are below this line. The box plus whiskers include all data from the 2.5% percentile to the 97.5% percentile, i.e., 95% of all data points.

Descriptive statistics for specific cellular fluorescence for all samples of *B. subtilis* GFP reporter strain inoculated with either a stationary phase (situation A) or an exponentially growing pre-culture (situation B) are listed in the supplementary material and plotted in Figure 2. Corresponding box-plots are presented in Figure 3.



**Figure 2** Descriptive statistics describing specific cellular fluorescence computed from flow cytometry analysis of *B. subtilis* AE099 in cultures inoculated with either a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). (Top left) Mean (black squares) and median (white squares). (Top right) standard deviation (triangles). (Bottom left) coefficient of variation (dots). (Bottom right) interquartile range (triangle, down). Each symbol can be linked to a point in the growth curves shown above the box-plots in Figure 3.



**Figure 3.** Box-plot showing the distribution of specific cellular fluorescence for *B. subtilis* GFP reporter strain inoculated with either a stationary phase or an exponentially growing pre-culture. The results were obtained by flow cytometry analysis. The points in the growth curves shown above the box-plots indicate when the samples for flow cytometry were harvested. **A.** Specific cellular fluorescence measured in a culture inoculated with a stationary phase pre-culture. **B.** Specific cellular fluorescence measured in a culture inoculated with an exponentially growing pre-culture (balanced growth). Chemically defined medium was used for cultures and pre-cultures.

In Figure 2 it is seen that the range of mean values for situation B is covered by the range for mean values for situation A. In both situation A and B, the tendency is that the mean (and median) increases during exponential growth.

The mean values of the culture inoculated with the stationary phase pre-culture did initially decrease, after which an increase was observed during the remaining of lag-phase (see Figure 3, top left, A). The initially decrease in mean specific cellular fluorescence may be found in a decrease in mean values of fluorescence per cell. The decrease in single cell fluorescence might be caused by the stationary phase resting cell used for inoculation contained mostly matured GFP, once the cells start to grow and producing GFP, the matured GFP was diluted, but the maturation of the newly synthesized GFP created a time frame where the overall fluorescence decreased. Once the newly produced GFP matured the fluorescence intensity of the cells

increased, and as exponential phase was reached a balanced between GFP expression/maturation and cell division rate the more robust part of the data indicated a steady state.

In situation B the sample means related to samples from 155 to 458 minutes are centered around a straight line with positive slope. The variation around the line is relatively small. The fact that the systematic pattern (the line) is very clear compared to the relatively small variation of means around the line indicates that an internal factor is causing this change. A similar tendency, although with larger variation (fluctuation) around the line, is observed for the exponential phase in A and strengthens the conclusion.

The immediate impression from the box-plots in Figure 3 is that the variation (interquartile ranges as well as the length of the whiskers) is larger *within* individual samples from situation A compared to *within* individual samples from situation B, indicating that an exponentially growing pre-culture lead to samples with smaller *within* sample variation.

Using Figure 2 to compare how the variation (SD, CV and IQR) changes *between* samples it is seen that the variation is relatively constant in situation B compared to situation A. From the box-plots it is seen that the sample distributions are generally skewed (longer whiskers above the box compared to below the box). As mentioned before, specifically for a skewed distribution, it can be relevant to pay attention to CV. In Figure 2 it is seen that CV is decreasing (roughly) with a constant rate in situation A. The last three values for CV in situation A are comparable to the values in situation B.

It should be noted that the measures of variation in Figure 2 (SD, CV and IQR) as well as the box-plots in Figure 3 do not reveal subpopulations in case they are present. This means that two samples may appear to be similar with regard to degree of heterogeneity, however one sample may consist of subpopulations and the other not. An investigation of possible subpopulations is made below using histograms as well as mixture distribution.

The conclusion based on Figure 2 and 3 is that after inoculation both the *within* sample heterogeneity (SD, CV and IRQ) as well as the fluctuation (variation) in the parameters in Figure 2 *between* samples is larger when the stationary phase pre-culture was used as inoculum compared to the used of an exponentially growing pre-culture.

#### **11.1.4.4 Histograms and investigating subpopulations by fitting mixture distributions**

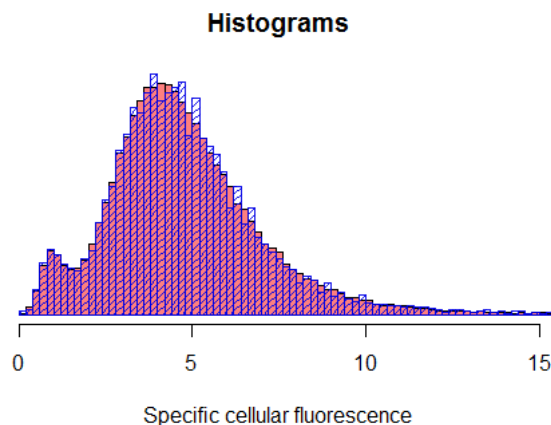
After calculating descriptive statistics and visualizing these using e.g. box-plots the natural next step in the analysis of heterogeneity is to investigate the shape of the histogram for the individual samples in more detail. The histograms for the samples in Figure 2 and 3 are presented in the supplementary material.

Two subpopulations are clearly seen in the histograms for the culture inoculated with the stationary phase pre-culture. The subpopulations are a main population wherefrom a considerably smaller subpopulation emerges and travels to the left in the histograms before it disappears. Except for 250 minutes after inoculation subpopulations are not visible in the histograms for samples from the exponentially growing pre-culture.

As mentioned in the introduction the skew-Laplace as well as the log-skew-Laplace probability distribution has received focus in relation to modelling the distribution of flow cytometry data. These distributions can be viewed as the overall distribution for a mixture of respectively normal distributions and log-normal distributions. However, for the general non-statistical practitioner, it can be a hurdle to work with these distributions. The statistical point-and-click software, JMP, offers a solution to overcome this hurdle: One of the options under the Distribution Platform in JMP is to fit mixtures of normal distributions. Fitting a mixture of normal distributions corresponds from a practical perspective to the skew-Laplace situation. If the mixtures of normal distributions it fitted to the log of the data, e.g.  $\log(\text{specific cellular fluorescence})$ , it similarly corresponds to the log-skew-Laplace situation.

For both situation A and B all continuous probability functions included in JMP was fitted to respectively specific cellular fluorescence and the log-transformed specific cellular fluorescence. Using Diagnostic Plots (probability plots) in JMP it was found that both in situation A and B it was possible to obtain a better fit, if specific cellular fluorescence were log-transformed. For the log-transformed data the continuous probability distributions fitted to each sample were compared using the objective AICc criteria (corrected Akaike Information Criteria). AICc is a measure of the relative quality of the fit of the distribution: the smaller the value of AICc the better the fit of the distribution. Generally the mixture of 3 normal distributions gave the best fit to the log-transformed data, and the mixture of 2 normal distributions gave the second best fit.

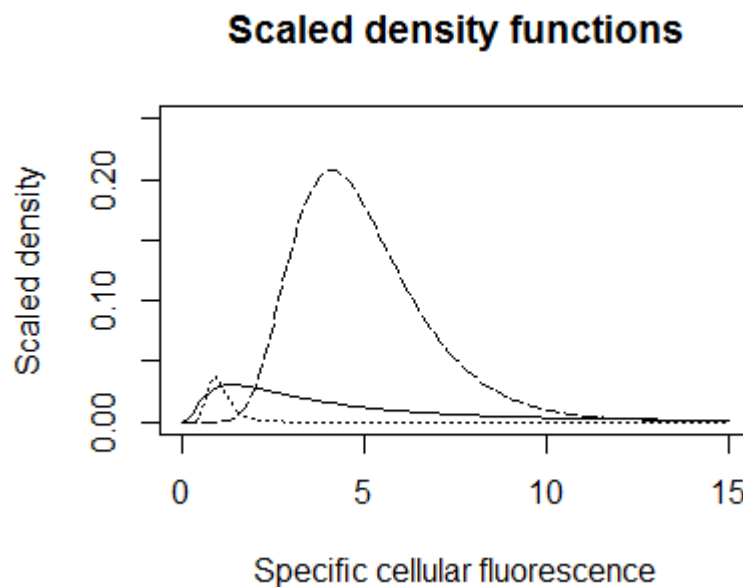
As an example the histogram for situation A,  $t=267$  minutes is shown using grey shading in Figure 4.



**Figure 4. Histograms for sample corresponding to stationary phase pre-culture,  $t=267$  minutes. The histogram for the actual data for specific cellular fluorescence is shown using blue shading. The histogram for data simulated based on the mixture of 3 normal distributions found using JMP is shown in the pink histogram behind the blue shaded histogram.**

The blue shaded histogram has minor deviations from the corresponding histogram in the supplementary material. The deviation is due to the actual binning of the data.

Using JMP it was found that with probabilities of respectively 0.15, 0.03 and 0.82 the log-transformed data origins from the three normal distributions:  $N(1.20, 0.93^2)$ ,  $N(-0.02, 0.31^2)$  and  $N(1.54, 0.36^2)$ . The theoretical probability density functions for these 3 distributions are shown in Figure 5. Note that in order to be able to compare Figure 4 and 5 directly, the x-axis in Figure 5 is specific cellular fluorescence and not the log-transformed data, and further the functions are scaled so that the total area under the 3 curves is equal to 1. The functions are scaled according to the probabilities found by JMP (i.e. 0.15, 0.03 and 0.82 as mentioned above).



**Figure 5. Scaled density functions for each of the three normal distribution in the mixture distribution fitted to  $\log(\text{specific cellular fluorescence})$  at  $t=267$  minutes in situation A. According to the mixture distribution 3% of the observations follow a  $N(-0.02, 0.31^2)$ -distribution (dotted line), 15% follow a  $N(1.20, 0.93^2)$ -distribution (full line) and 82% follow a  $N(1.54, 0.36^2)$ -distribution (dashed line). The density functions are scaled using the respectively percentages so that the total area under the three curves is 1.**

In order to check how well the mixture distribution fitted to the data, an equal number of observations ( $N=102292$ ) were simulated from the mixture distribution and plotted in a histogram in Figure 4 together with the histogram for the real data. The simulated data are shown with the white histogram behind the shaded grey histogram. The histograms are for practical purposes identical.

The practical interpretation of the fit of the mixture distribution is that the sample at  $t=267$  minutes consist of three subpopulations. However, it should be kept in mind that the subpopulations found using JMP may or may not correspond to real physical subpopulations in the sample. The practitioner should critically interpret the results regarding subpopulations found by JMP while taking biological knowledge into account. In the example above maybe a guess would be that the distribution corresponding to 3% of the data is the result of

noise or of the two other distributions deviating some from the log-normal distribution (if the log-transformed data follows a normal distribution, the original data is log-normally distributed).

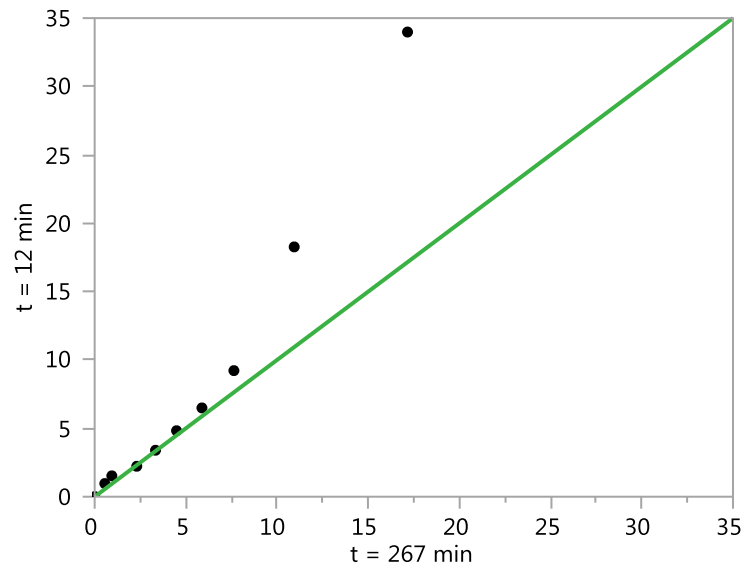
Even though all subpopulations identified by fitting mixture distributions in JMP may not correspond to real, physical subpopulations, the information from JMP provides valuable extra input regarding how subpopulations might contribute to the heterogeneity of the individual sample.

An alternative to identify and model the subpopulations seen in the data in Figure 4 would be to use gating to separate data in two groups and thereafter fit a distribution to each group separately. It is seen from figure 4, though, that using a cut-value of specific cellular fluorescence to cut the distribution in two groups would be a very rough way to separate the two groups – some of the values below the cut-value would actually belong to the group of larger values and vice versa. These observations would be misclassified if the gating technique is used in this example.

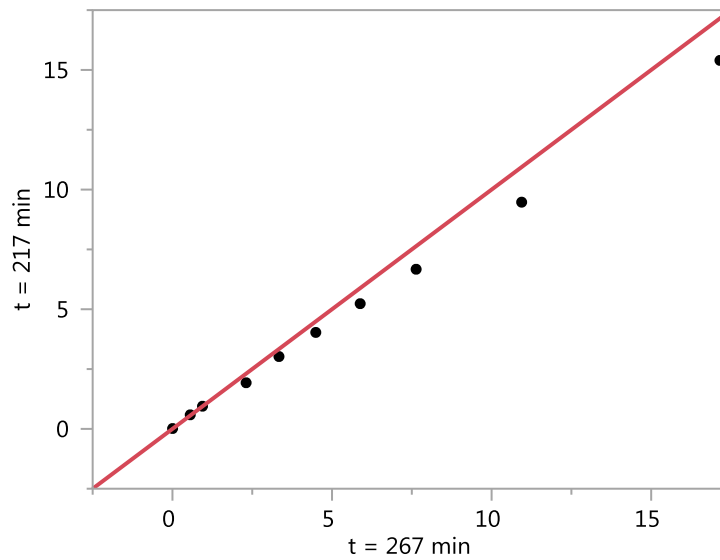
JMP indicates that also in situation B the log-transformed data is generally best fitted by a mixture of three normal distributions. However, in situation B subpopulations are not immediately visible from the histograms more investigations need to be done in relation to investigating to which degree subpopulations identified by JMP are

#### **11.1.4.5 Comparing distributions of different samples**

A QQ-plot can be used as a graphical tool to visualize the similarity or heterogeneity between distributions from two samples. In the supplementary material a table of a number of quantiles is listed for each sample. If two (empirical distributions) are identical, their quantiles will be identical too. Thus, plotting matching quantiles against each other (i.e. the 5% quantile from one distribution against the 5% quantile from the other distribution; the 10% quantile from one distribution against the 10% quantile from the other etc.) will form a straight line with intercept = 0 and slope = 1. If quantiles from two non-identical distributions are plotted against each other, the points will not form a straight line (with intercept = 0 and slope = 1). The deviation from the straight line will reflect the dissimilarity between the distributions. As an example the quantiles in the supplementary material for the sample in situation A,  $t = 12$  minutes are plotted against the quantiles for the sample in situation A,  $t = 267$  minutes in Figure 6. Similarly, the quantiles from the sample at  $t = 217$  are plotted against the quantiles from the sample at  $t = 267$  minutes in Figure 7. It is seen that the points in Figure 7 are closer to the line with intercept = 0 and slope = 1 compared to the points in Figure 6. This shows that the samples at  $t = 217$  minutes and  $t = 267$  minutes are more similar with regard to distribution. In Figure 6 and 7 the 100% quantiles have not been included. The reason is that these values are relatively large and are making it difficult to see the pattern in the points in the rest of the plot.



**Figure 6.** Quantiles from the samples at  $t = 12$  minutes and  $t = 267$  minutes (both from situation A) plotted against each other. If the two empirical distributions (histograms) are identical, the points will form a straight line with intercept = 0 and slope = 1. The points reflecting both the small as well as the large quantiles are above the line, showing that the distribution at  $t = 267$  minutes has more mass at the small as well as at the large values of specified cellular fluorescence (see the histograms in the supplementary material) compared to the distribution at  $t = 12$  minutes.



**Figure 7.** Quantiles from the samples at  $t = 217$  minutes and  $t = 267$  minutes (both from situation A) plotted against each other. If the two empirical distributions (histograms) are identical, the points will form a straight line with intercept = 0 and slope = 1. The fact that the points are lying below the line for the larger quantiles show that the histogram for  $t = 267$  minutes has a longer tail to the right compared to the histogram for  $t = 217$  minutes. This can be difficult to see when comparing the histograms in the supplementary material directly, but the qq-plot makes it clear.

### 11.1.5 Discussion

A main challenge in relation to assessing heterogeneity of flow cytometry data is that there is no definition of what precisely is meant by the vague, general term: heterogeneity. The lack of definition of what is meant by heterogeneity makes it difficult objectively to assess and discuss heterogeneity of flow cytometry data.

In this article we have introduced the concept of splitting heterogeneity into heterogeneity *within* a sample of flow cytometry data versus heterogeneity *between* samples.

When we look at an individual sample both the overall variation as well as the shape of the distribution (histogram) and possible existence of subpopulations are relevant. The location (mean and median) may be of less relevance in relation to assessing heterogeneity when only one sample is considered. Currently there is no common guidance on how to put all these aspects of *within* sample heterogeneity into one overall measure of heterogeneity, an index of heterogeneity – and it is questionable if it would be meaningful to try to do so.

When comparing the heterogeneity across samples it is relevant to investigate how all these aspects of heterogeneity (overall variation, shape of distribution as well as subpopulations) as well as how the location (mean or median) varies *between* the samples.

Besides the input to a terminology related to assessing heterogeneity of flow cytometry data, this article also describes how heterogeneity as defined/described by the suggested terminology can be quantified using



descriptive statistics, box-plots, mixture distributions and qq-plots. All of these statistical methods are easily applied using e.g. the statistical point-and-click software JMP®.

A typical statistical analysis would in many practical applications include a more formal statistical hypothesis test. A relevant hypothesis in relation to flow cytometry data is: is the variation (here variance) within samples equal for samples taken at different time points and/or different experimental conditions, or do the variation within these groups differ? Or similarly: are the mean of different samples (and/or experimental conditions) equal or do they differ? However, performing standard tests for homogeneity of variance and e.g. ANOVA or mixed models (testing if means are equal) is not suitable in the case of flow cytometry data as for this practical application the number of observations in each sample can be in the order 100.000. With this amount of observations even a very small difference between the means or a very small inhomogeneity between variances will be declared statistically significant by the test procedure. One general way to overcome the problem of (too) large sample sizes in relation to testing is to use tests for equivalence of means and/or variances (Frey, 2010; Wellek, 2003), instead of the more standard tests mentioned previously. In general, equivalence testing is used when the goal is not to show that two or more parameters differ, but rather to show that they are roughly the same. Tests for equivalence were considered in relation to the analysis of the data presented in this article. However, a requirement for equivalence testing is that a difference between means (or ratio of variances) can be specified, which from a practical perspective can be considered as a limit for means or variances being 'roughly the same'. Setting up such specifications is not feasible for flow cytometry data like specific cellular fluorescence.

As an alternative to the statistical tests mentioned above, non-parametric statistical tests can be considered. Non-parametric statistical tests are based on no assumptions about the probability distribution of the relevant parameter. However, with the data presented here, the large-sample problems mentioned above is also a problem for non-parametrical, distribution free tests.

In previous work that also focus on using statistical methods to describe heterogeneity of cellular properties in isogenic cultures a number of probability distributions have been investigated. Some of these can be interpreted as modelling the overall distribution of a mixture of subpopulations. So far it has not been possible to find a common family of distributions that fit well in all situations of flow cytometry data.

In this paper a tool to fit distributions to subpopulations individually (in contrast to fitting a distribution to the overall mixture of the subpopulations) has been presented. Knowledge of the distribution including location of the individual subpopulations adds valuable knowledge to understanding the heterogeneity of flow cytometry data.

## **Acknowledgements**

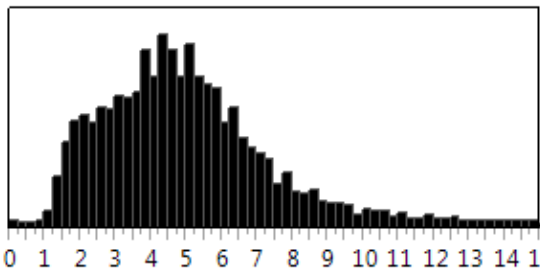
This work was supported by the Danish Council for Strategic Research. We thank Inka Sastalla for the kind gift of plasmid pSW4-GFPopt. We thank Flemming Hansen for fluorescent proteins guidance.

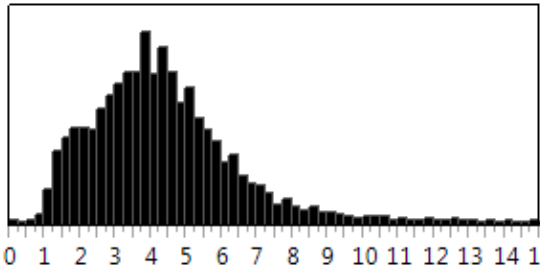
### 11.1.6 References

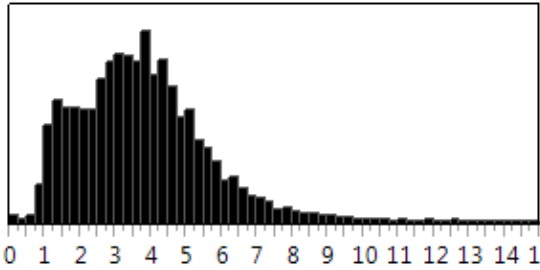
- Bagnold, R. a. & Barndorff-Nielsen, O. (1980).** The pattern of natural size distributions. *Sedimentology* **27**, 199–207.
- Barndorff-Nielsen, O. (1977).** Exponentially decreasing distributions for the logarithm of particle Size. *Proc R Soc London A* **353**, 401–419.
- Fieller, N. R. J., Flenley, E. C. & Olbricht, W. (1992).** Statistics of Particle Size Data. *R Stat Soc* **41**, 127–146.
- Frey, J. (2010).** Testing for equivalence of variances using Hartley’s ratio. *Can J Stat* **38**, 647–664.
- Julià, O. & Vives-Rego, J. (2005).** Skew-Laplace distribution in Gram-negative bacterial axenic cultures: New insights into intrinsic cellular heterogeneity. *Microbiology* **151**, 749–755.
- Julià, O., Vidal-Mas, J., Panikov, N. S. & Vives-Rego, J. (2010).** Skew-laplace and cell-size distribution in microbial axenic cultures: statistical assessment and biological interpretation. *Int J Microbiol* **2010**, 191585.
- Kotz, S., Kozubowski, T. J. & Podgórski, K. (2001).** *The laplace distribution and generalizations*. Birkhäuser Boston.
- Pedersen, A. E., Thyregod, C., Kilstrup, M. & Martinussen, J. (2015).** Heterogeneity at a single cell level in homogenous cultures. *in preparation*.
- Prats, C., Giró, A., Ferrer, J., López, D. & Vives-Rego, J. (2008).** Analysis and IbM simulation of the stages in bacterial lag phase: Basis for an updated definition. *J Theor Biol* **252**, 56–68.
- Prats, C., Ferrer, J., López, D., Giró, A. & Vives-rego, J. (2010).** On the evolution of cell size distribution during bacterial growth cycle : Experimental observations and individual-based model simulations. *J Microbiol* **4**, 400–407.
- Saxild, H. H. & Nygaard, P. (1987).** Genetic and physiological characterization of *Bacillus subtilis* mutants resistant to purine analogs. *J Bacteriol* **169**, 2977–2983.
- Wellek, S. (2003).** *Testing statistical hypotheses of equivalence*. Chapman and Hall, New York.

### 11.1.7 Supplementary material

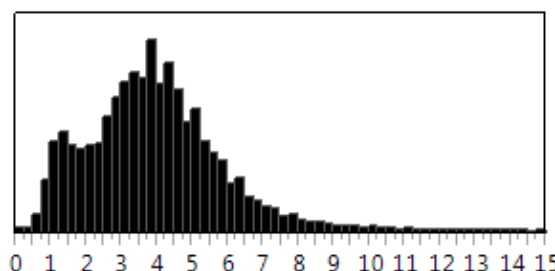
Table 1 JMP® analysis of specific cellular fluorescence measured for the *B. subtilis* reporter strain cultivated in chemically defined medium inoculated with a stationary phase pre-culture.

Specific cellular fluorecence, t = 12 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	125,215	Mean	5,8014258
	99.5%		33,9821	Std Dev	5,0136402
	97.5%		18,2692	N	99102
	90.0%		9,2224	CV	86,420828
	75.0%	quartile	6,49382	Median	4,8260715
	50.0%	median	4,82607	Interquartile Range	3,095608
	25.0%	quartile	3,39821		
	10.0%		2,22667		
	2.5%		1,52614		
	0.5%		0,96466		
	0.0%	minimum	0,00743		

Specific cellular fluorecence, t = 57 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	122,983	Mean	4,9404978
	99.5%		28,9026	Std Dev	4,2304848
	97.5%		15,3993	N	91597
	90.0%		7,5667	CV	85,628715
	75.0%	quartile	5,52316	Median	4,1792057
	50.0%	median	4,17921	Interquartile Range	2,4999454
	25.0%	quartile	3,02321		
	10.0%		1,98096		
	2.5%		1,34497		
	0.5%		0,94746		
	0.0%	minimum	0,00806		

Specific cellular fluorecence, t = 112 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	126,346	Mean	4,1433907
	99.5%		22,4679	Std Dev	3,3358926
	97.5%		11,04	N	97535
	90.0%		6,43567	CV	80,511178
	75.0%	quartile	4,91367	Median	3,6847349
	50.0%	median	3,68473	Interquartile Range	2,4108076
	25.0%	quartile	2,50287		
	10.0%		1,53993		
	2.5%		1,046		
	0.5%		0,64938		
	0.0%	minimum	0,02595		

### Specific cellular fluorescence, t = 159 min



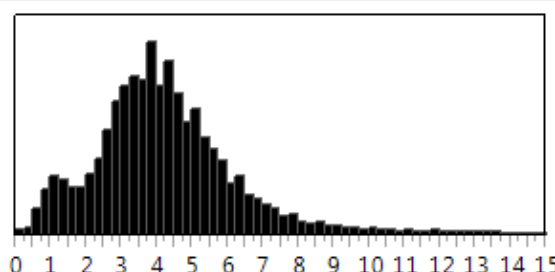
#### Quantiles

100.0%	maximum	136,999
99.5%		17,9435
97.5%		9,64662
90.0%		6,43567
75.0%	quartile	5,04807
50.0%	median	3,85423
25.0%	quartile	2,7139
10.0%		1,53993
2.5%		0,97338
0.5%		0,61507
0.0%	minimum	0,01358

#### Summary Statistics

Mean	4,1584118
Std Dev	2,8190522
N	101192
CV	67,791558
Median	3,8542289
Interquartile Range	2,3341663

### Specific cellular fluorescence, t = 217 min



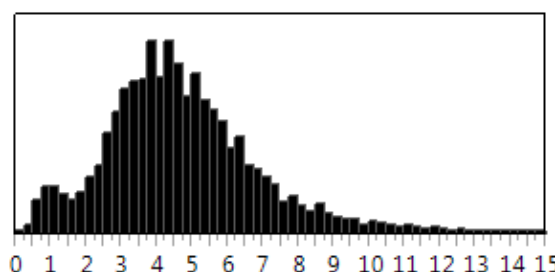
#### Quantiles

100.0%	maximum	95,6024
99.5%		15,3993
97.5%		9,47464
90.0%		6,67143
75.0%	quartile	5,23299
50.0%	median	4,03152
25.0%	quartile	3,02321
10.0%		1,92822
2.5%		0,94746
0.5%		0,58821
0.0%	minimum	0,01582

#### Summary Statistics

Mean	4,328751
Std Dev	2,4509804
N	102037
CV	56,620961
Median	4,0315194
Interquartile Range	2,2097781

### Specific cellular fluorescence, t = 267 min



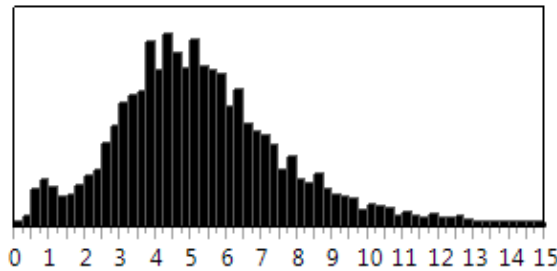
#### Quantiles

100.0%	maximum	94,7464
99.5%		17,1544
97.5%		10,9411
90.0%		7,63506
75.0%	quartile	5,88208
50.0%	median	4,49101
25.0%	quartile	3,33762
10.0%		2,30824
2.5%		0,94174
0.5%		0,55731
0.0%	minimum	0,00457

#### Summary Statistics

Mean	4,8791324
Std Dev	2,702958
N	102292
CV	55,398332
Median	4,4910072
Interquartile Range	2,5444597

### Specific cellular fluorescence, t = 330 min



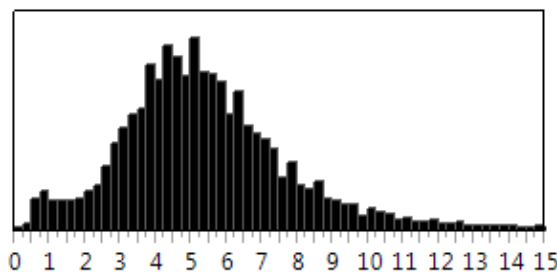
#### Quantiles

100.0%	maximum	83,5363
99.5%		20,3514
97.5%		12,7488
90.0%		8,73788
75.0%	quartile	6,67143
50.0%	median	5,09368
25.0%	quartile	3,75162
10.0%		2,59455
2.5%		0,98217
0.5%		0,55731
0.0%	minimum	0,00914

#### Summary Statistics

Mean	5,5434032
Std Dev	3,1049093
N	101818
CV	56,010888
Median	5,0936752
Interquartile Range	2,919808

### Specific cellular fluorescence, t = 364 min



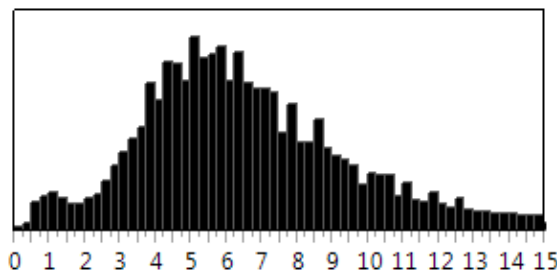
#### Quantiles

100.0%	maximum	79,1476
99.5%		18,9384
97.5%		12,1881
90.0%		8,5821
75.0%	quartile	6,67143
50.0%	median	5,1397
25.0%	quartile	3,88905
10.0%		2,76316
2.5%		1,08432
0.5%		0,59889
0.0%	minimum	0,01731

#### Summary Statistics

Mean	5,5639397
Std Dev	2,9278738
N	101273
CV	52,622313
Median	5,1396968
Interquartile Range	2,7823752

### Specific cellular fluorescence, t = 397 min



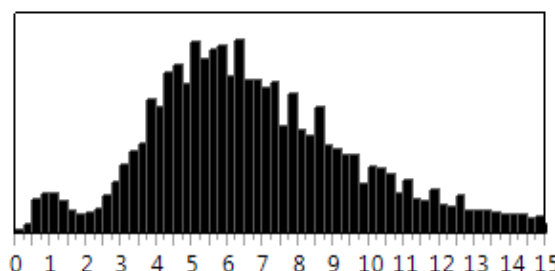
#### Quantiles

100.0%	maximum	121,881
99.5%		31,059
97.5%		18,9384
90.0%		11,9709
75.0%	quartile	8,65964
50.0%	median	6,32093
25.0%	quartile	4,61384
10.0%		3,33762
2.5%		1,34557
0.5%		0,66714
0.0%	minimum	0,01358

#### Summary Statistics

Mean	7,2820722
Std Dev	4,6903188
N	100875
CV	64,409122
Median	6,3209339
Interquartile Range	4,0458036

### Specific cellular fluorescence, t = 430 min



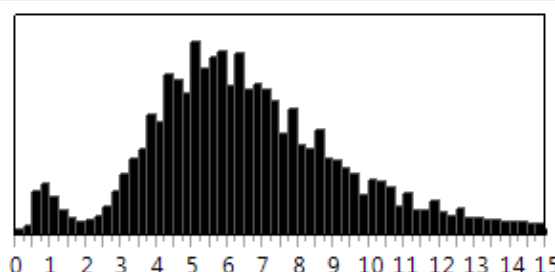
#### Quantiles

100.0%	maximum	94,7464
99.5%		32,1968
97.5%		19,4564
90.0%		12,2983
75.0%	quartile	8,97687
50.0%	median	6,55249
25.0%	quartile	4,86968
10.0%		3,55452
2.5%		1,26346
0.5%		0,6378
0.0%	minimum	0,01963

#### Summary Statistics

Mean	7,5562828
Std Dev	4,8257959
N	100762
CV	63,864681
Median	6,5524882
Interquartile Range	4,1071961

### Specific cellular fluorescence, t = 483 min



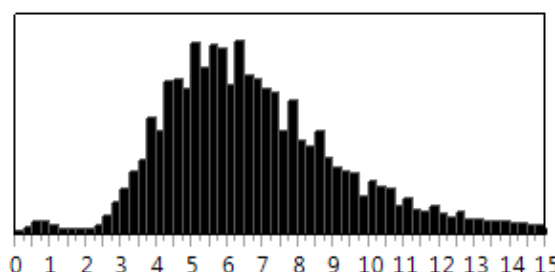
#### Quantiles

100.0%	maximum	95,6024
99.5%		28,6438
97.5%		17,0008
90.0%		11,2404
75.0%	quartile	8,4291
50.0%	median	6,32093
25.0%	quartile	4,78286
10.0%		3,52269
2.5%		1,01815
0.5%		0,58138
0.0%	minimum	0,00422

#### Summary Statistics

Mean	7,0993327
Std Dev	4,249584
N	100739
CV	59,858921
Median	6,3209339
Interquartile Range	3,6462427

### Specific cellular fluorescence, t = 554 min



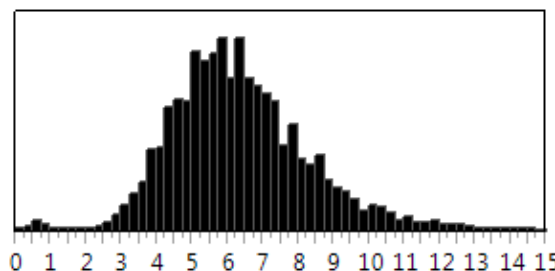
#### Quantiles

100.0%	maximum	147,221
99.5%		25,7132
97.5%		15,6788
90.0%		10,7461
75.0%	quartile	8,35363
50.0%	median	6,49382
25.0%	quartile	5,04807
10.0%		3,99542
2.5%		2,91638
0.5%		0,75667
0.0%	minimum	0,00777

#### Summary Statistics

Mean	7,1613258
Std Dev	3,732217
N	99890
CV	52,116286
Median	6,4938163
Interquartile Range	3,3055598

### Specific cellular fluorescence, t = 632 min



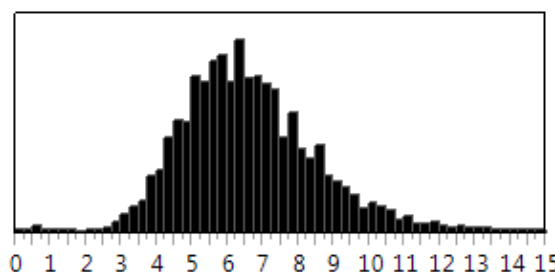
#### Quantiles

100.0%	maximum	62,6434
99.5%		15,1247
97.5%		11,2404
90.0%		8,97687
75.0%	quartile	7,49894
50.0%	median	6,15265
25.0%	quartile	5,09368
10.0%		4,17921
2.5%		3,24877
0.5%		0,76965
0.0%	minimum	0,0075

#### Summary Statistics

Mean	6,4429573
Std Dev	2,201229
N	99777
CV	34,164886
Median	6,1526541
Interquartile Range	2,4052669

### Specific cellular fluorescence, t = 692 min



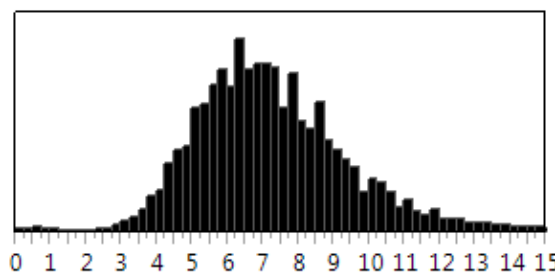
#### Quantiles

100.0%	maximum	78,4389
99.5%		14,9893
97.5%		11,3419
90.0%		9,13982
75.0%	quartile	7,77365
50.0%	median	6,43567
25.0%	quartile	5,32798
10.0%		4,45079
2.5%		3,49115
0.5%		1,54514
0.0%	minimum	0,03162

#### Summary Statistics

Mean	6,6977687
Std Dev	2,1591298
N	99874
CV	32,236554
Median	6,4356698
Interquartile Range	2,4456714

### Specific cellular fluorescence, t = 762 min



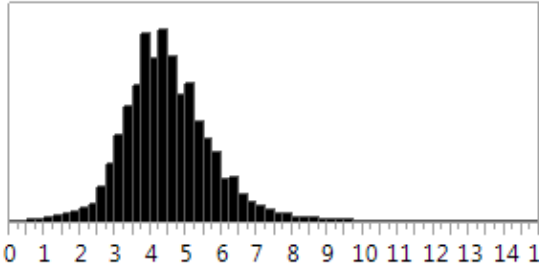
#### Quantiles

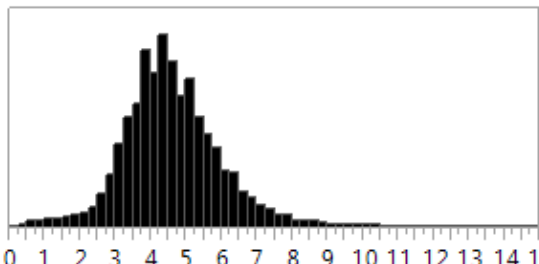
100.0%	maximum	71,6917
99.5%		16,6977
97.5%		12,6346
90.0%		10,0904
75.0%	quartile	8,50526
50.0%	median	7,04136
25.0%	quartile	5,82942
10.0%		4,86968
2.5%		3,88905
0.5%		2,64165
0.0%	minimum	0,00626

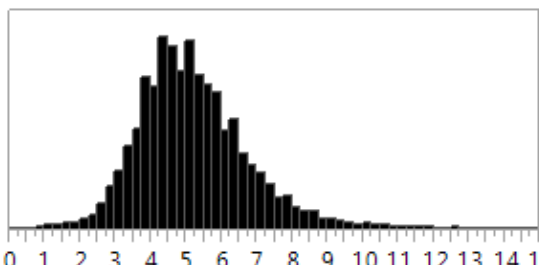
#### Summary Statistics

Mean	7,352805
Std Dev	2,3942887
N	99832
CV	32,56293
Median	7,0413552
Interquartile Range	2,6758428

Table 7 JMP® analysis of specific cellular fluorescence measured for the *B. subtilis* reporter strain cultivated in chemically defined medium inoculated with an exponentially growing pre-culture (a balanced growing culture)

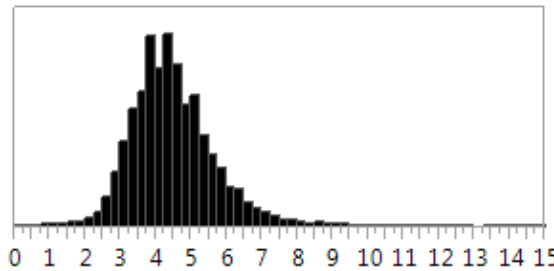
Specific cellular fluorescence, t = 31 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	55,2316	Mean	4,5154153
	99.5%		8,89649	Std Dev	1,2646213
	97.5%		7,23394	N	98994
	90.0%		6,04296	CV	28,006754
	75.0%	quartile	5,18613	Median	4,410941
	50.0%	median	4,41094	Interquartile Range	1,4681075
	25.0%	quartile	3,71803		
	10.0%		3,13396		
	2.5%		2,35014		
	0.5%		1,36999		
	0.0%	minimum	0,12409		

Specific cellular fluorescence, t = 65 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	44,5079	Mean	4,6224173
	99.5%		9,64662	Std Dev	1,4399735
	97.5%		7,70404	N	98279
	90.0%		6,32093	CV	31,151958
	75.0%	quartile	5,37612	Median	4,4910072
	50.0%	median	4,49101	Interquartile Range	1,6244983
	25.0%	quartile	3,75162		
	10.0%		3,1059		
	2.5%		1,96322		
	0.5%		0,82047		
	0.0%	minimum	0,01486		

Specific cellular fluorescence, t = 98 min					
	Quantiles		Summary Statistics		
	100.0%	maximum	99,1046	Mean	5,2536539
	99.5%		11,7574	Std Dev	1,7535393
	97.5%		8,97687	N	99518
	90.0%		7,23394	CV	33,377519
	75.0%	quartile	6,04296	Median	5,0480657
	50.0%	median	5,04807	Interquartile Range	1,8637582
	25.0%	quartile	4,17921		
	10.0%		3,49115		
	2.5%		2,7139		
	0.5%		1,65482		
	0.0%	minimum	0,07499		



### Specific cellular fluorescence, t = 124 min



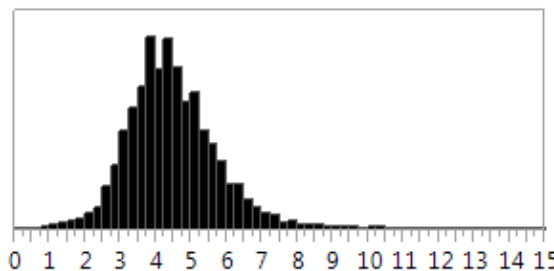
#### Quantiles

100.0%	maximum	28,6438
99.5%		8,81051
97.5%		7,16917
90.0%		5,93523
75.0%	quartile	5,1397
50.0%	median	4,37144
25.0%	quartile	3,71803
10.0%		3,21968
2.5%		2,61799
0.5%		1,77956
0.0%	minimum	0,12079

#### Summary Statistics

Mean	4,5022725
Std Dev	1,1864082
N	99815
CV	26,35132
Median	4,3714448
Interquartile Range	1,4216701

### Specific cellular fluorescence, t = 155 min



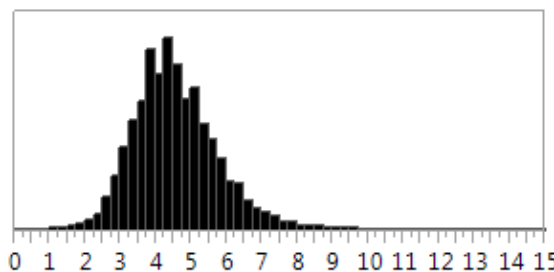
#### Quantiles

100.0%	maximum	46,1384
99.5%		9,3898
97.5%		7,36525
90.0%		6,04296
75.0%	quartile	5,18613
50.0%	median	4,37144
25.0%	quartile	3,68473
10.0%		3,07809
2.5%		2,30824
0.5%		1,38237
0.0%	minimum	0,05882

#### Summary Statistics

Mean	4,5033407
Std Dev	1,330517
N	99260
CV	29,545111
Median	4,3714448
Interquartile Range	1,5013993

### Specific cellular fluorescence, t = 186 min



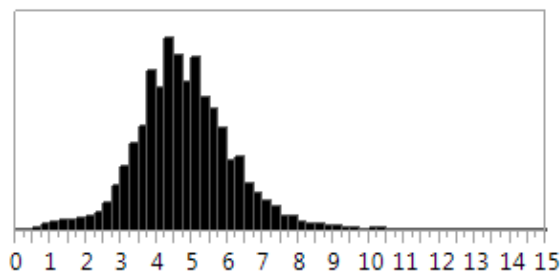
#### Quantiles

100.0%	maximum	79,1476
99.5%		9,47464
97.5%		7,36525
90.0%		6,09756
75.0%	quartile	5,28027
50.0%	median	4,45079
25.0%	quartile	3,78552
10.0%		3,21968
2.5%		2,57132
0.5%		1,77828
0.0%	minimum	0,03523

#### Summary Statistics

Mean	4,6049574
Std Dev	1,2963794
N	241797
CV	28,151821
Median	4,4507941
Interquartile Range	1,4947562

### Specific cellular fluorescence, t = 220 min



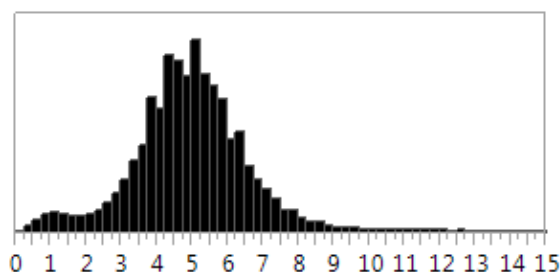
#### Quantiles

100.0%	maximum	65,5249
99.5%		9,3898
97.5%		7,77365
90.0%		6,49382
75.0%	quartile	5,57306
50.0%	median	4,74003
25.0%	quartile	3,95964
10.0%		3,24877
2.5%		2,10969
0.5%		1,06499
0.0%	minimum	0,07234

#### Summary Statistics

Mean	4,8077441
Std Dev	1,4128172
N	246474
CV	29,386281
Median	4,7400317
Interquartile Range	1,6134154

### Specific cellular fluorescence, t = 250 min



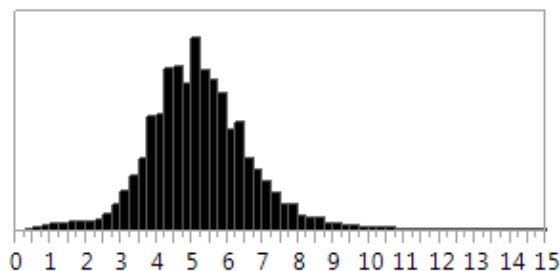
#### Quantiles

100.0%	maximum	37,5162
99.5%		11,5478
97.5%		8,27883
90.0%		6,79253
75.0%	quartile	5,88208
50.0%	median	4,95807
25.0%	quartile	4,06794
10.0%		3,13396
2.5%		1,30975
0.5%		0,65525
0.0%	minimum	0,18106

#### Summary Statistics

Mean	4,976986
Std Dev	1,6756418
N	98829
CV	33,667803
Median	4,9580682
Interquartile Range	1,8141401

### Specific cellular fluorescence, t = 283 min



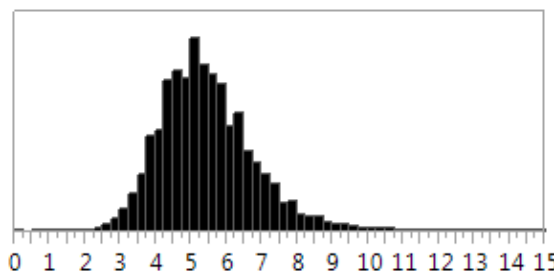
#### Quantiles

100.0%	maximum	44,9101
99.5%		9,91046
97.5%		8,2047
90.0%		6,97831
75.0%	quartile	6,04296
50.0%	median	5,1397
25.0%	quartile	4,3323
10.0%		3,61904
2.5%		2,52548
0.5%		1,18249
0.0%	minimum	0,29164

#### Summary Statistics

Mean	5,2194529
Std Dev	1,4415119
N	99726
CV	27,618064
Median	5,1396968
Interquartile Range	1,7106616

### Specific cellular fluorescence, t = 317 min



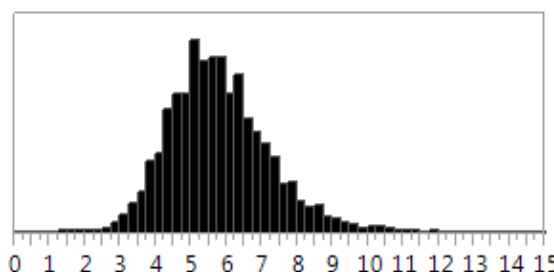
#### Quantiles

100.0%	maximum	68,539
99.5%		10,2735
97.5%		8,50526
90.0%		7,16917
75.0%	quartile	6,20824
50.0%	median	5,32798
25.0%	quartile	4,57253
10.0%		3,95964
2.5%		3,33762
0.5%		2,73842
0.0%	minimum	0,01472

#### Summary Statistics

Mean	5,4713781
Std Dev	1,3707193
N	99981
CV	25,052541
Median	5,3279789
Interquartile Range	1,6357169

### Specific cellular fluorescence, t = 356 min



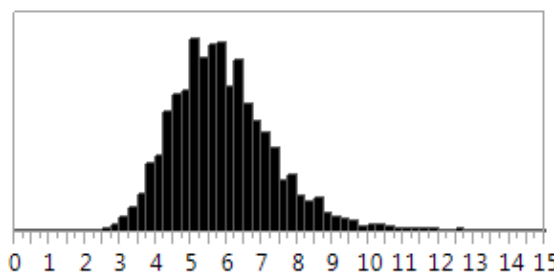
#### Quantiles

100.0%	maximum	35,2269
99.5%		10,9411
97.5%		9,13982
90.0%		7,63506
75.0%	quartile	6,67143
50.0%	median	5,67422
25.0%	quartile	4,86968
10.0%		4,17921
2.5%		3,42891
0.5%		2,64165
0.0%	minimum	0,01811

#### Summary Statistics

Mean	5,8332886
Std Dev	1,4747985
N	99962
CV	25,282453
Median	5,674221
Interquartile Range	1,8017519

### Specific cellular fluorescence, t = 394 min



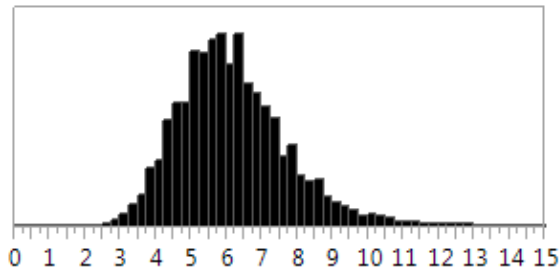
#### Quantiles

100.0%	maximum	62,0824
99.5%		11,1397
97.5%		9,2224
90.0%		7,70404
75.0%	quartile	6,7317
50.0%	median	5,77722
25.0%	quartile	4,95807
10.0%		4,25507
2.5%		3,61904
0.5%		3,07809
0.0%	minimum	0,07432

#### Summary Statistics

Mean	5,9329219
Std Dev	1,5026398
N	100037
CV	25,327146
Median	5,7772179
Interquartile Range	1,7736356

### Specific cellular fluorescence, t = 432 min



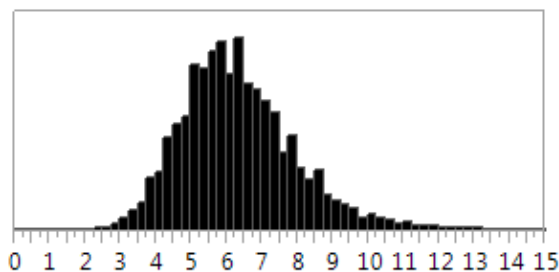
#### Quantiles

100.0%	maximum	87,3788
99.5%		11,8637
97.5%		9,82172
90.0%		8,2047
75.0%	quartile	7,04136
50.0%	median	6,04296
25.0%	quartile	5,1397
10.0%		4,41094
2.5%		3,65174
0.5%		3,07809
0.0%	minimum	0,03995

#### Summary Statistics

Mean	6,2047996
Std Dev	1,6602301
N	100041
CV	26,757191
Median	6,0429639
Interquartile Range	1,9016584

### Specific cellular fluorescence, t = 458 min



#### Quantiles

100.0%	maximum	44,5079
99.5%		12,5215
97.5%		10,1815
90.0%		8,4291
75.0%	quartile	7,2993
50.0%	median	6,15265
25.0%	quartile	5,23299
10.0%		4,49101
2.5%		3,68473
0.5%		3,05053
0.0%	minimum	0,06552

#### Summary Statistics

Mean	6,3770742
Std Dev	1,7086178
N	100212
CV	26,793131
Median	6,1526541
Interquartile Range	2,0663095

## 11.2 Supplementary materials - section 3

Box-plots and descriptive statistics for flow cytometry data (forward scatter per cell and fluorescence per cell) – supplementary materials for the manuscript “Heterogeneity at a single cell level in homogenous cultures”.

### 11.2.1 Heterogeneity determined in populations of *B. subtilis* growing at three different rates

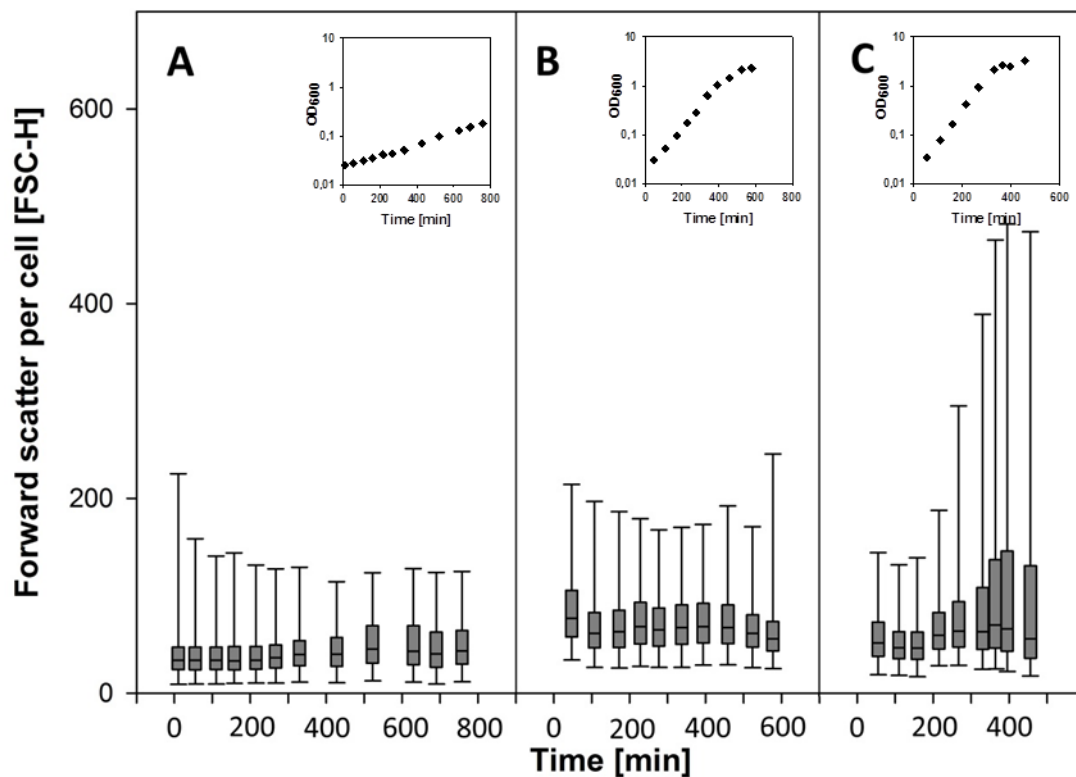
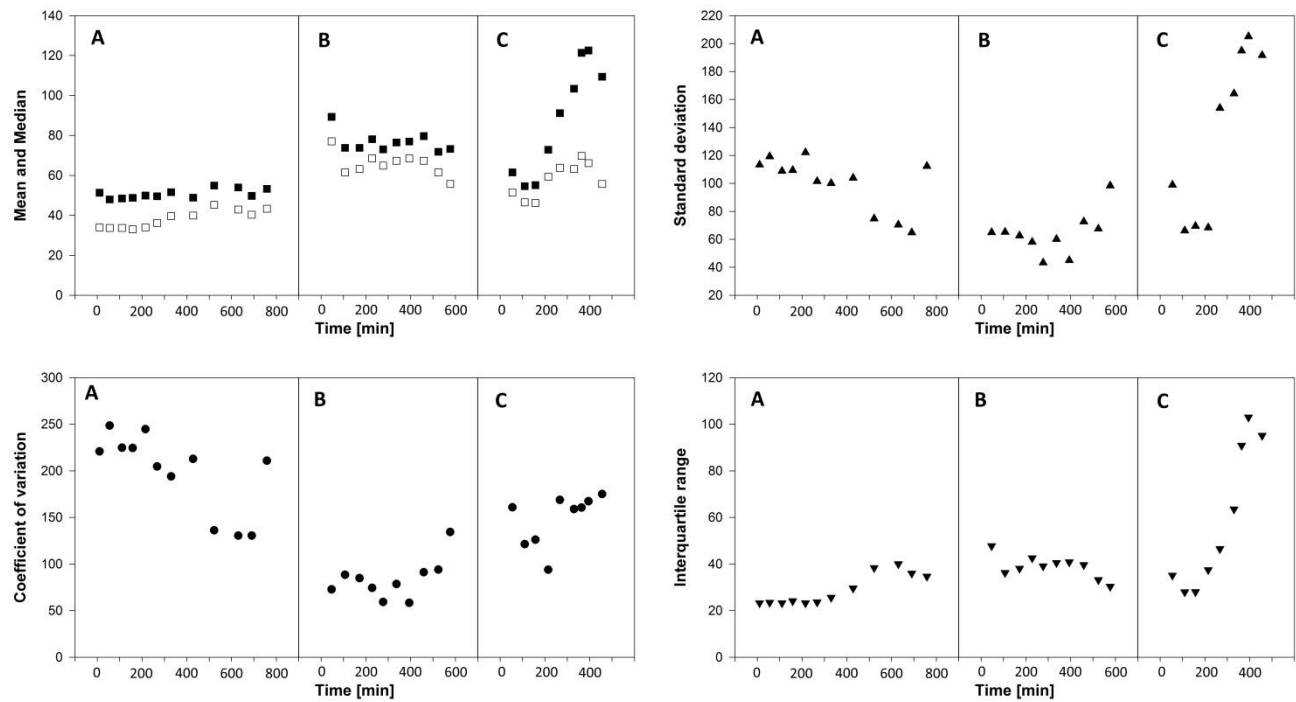


Figure 37 Box-plots showing the distribution of forward scatter per cell, for *B. subtilis* reporter strain AE099 cultivated as balanced cultures in chemically defined medium supplemented with (A) ribose (doubling time 256 min). (B) glucose (doubling time 67 min) and (C) glucose and glutamate (doubling time 45 min).



**Figure 38** Descriptive statistics based on forward scatter per cell, for *B. subtilis* reporter strain AE099 cultivated as balanced cultures in chemically defined medium supplemented with (A) ribose (doubling time 256 min). (B) glucose (doubling time 67 min) and (C) glucose and glutamate (doubling time 45 min). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 37.

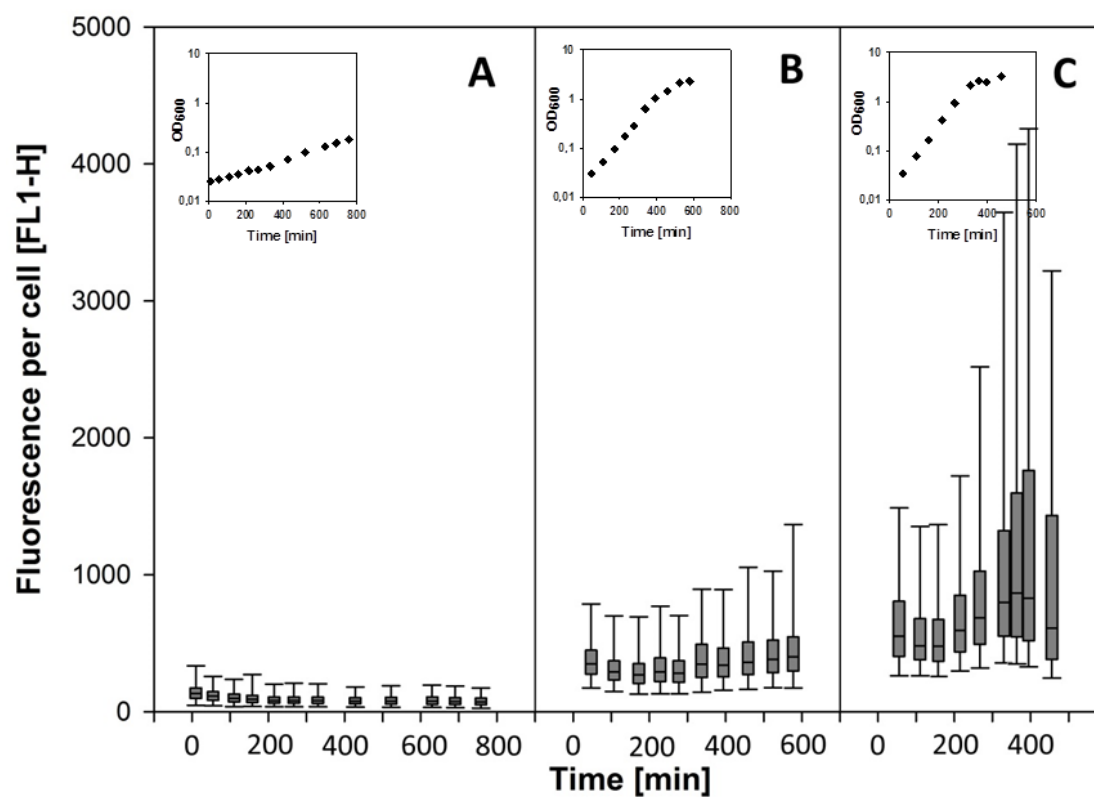
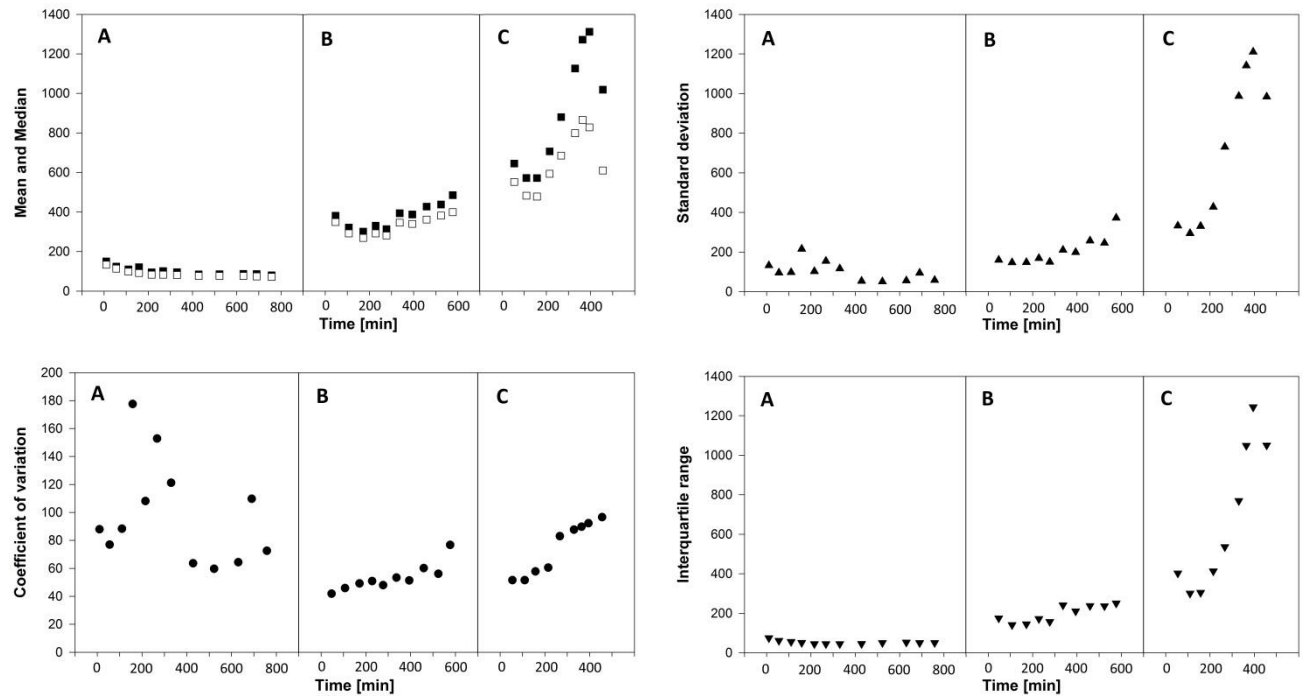


Figure 39 Box-plots showing the distribution of fluorescence per cell, for *B. subtilis* reporter strain AE099 cultivated as balanced cultures in chemically defined medium supplemented with (A) ribose (doubling time 256 min). (B) glucose (doubling time 67 min) and (C) glucose and glutamate (doubling time 45 min).



**Figure 40** Descriptive statistics based on fluorescence per cell for *B. subtilis* reporter strain AE099 cultivated as balanced cultures in chemically defined medium supplemented with (A) Ribose (doubling time 256 min). (B) Glucose (doubling time 67 min) and (C) glucose and glutamate (doubling time 45 min). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 39.



### 11.2.2 Heterogeneity determined in populations of *B. subtilis* growing in complex or chemically defined medium

The *B. subtilis* reporter strain was propagated in complex and chemically defined medium inoculated with a stationary phase complex medium inoculum or an exponential growing chemically defined inoculum.

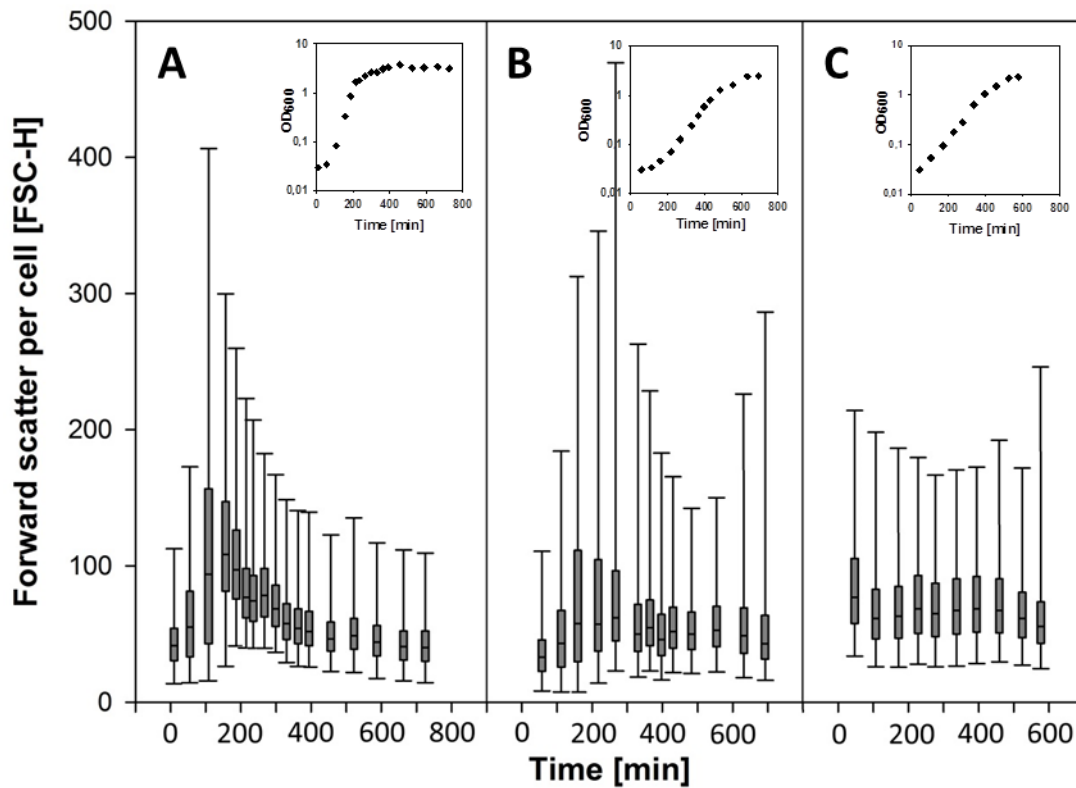
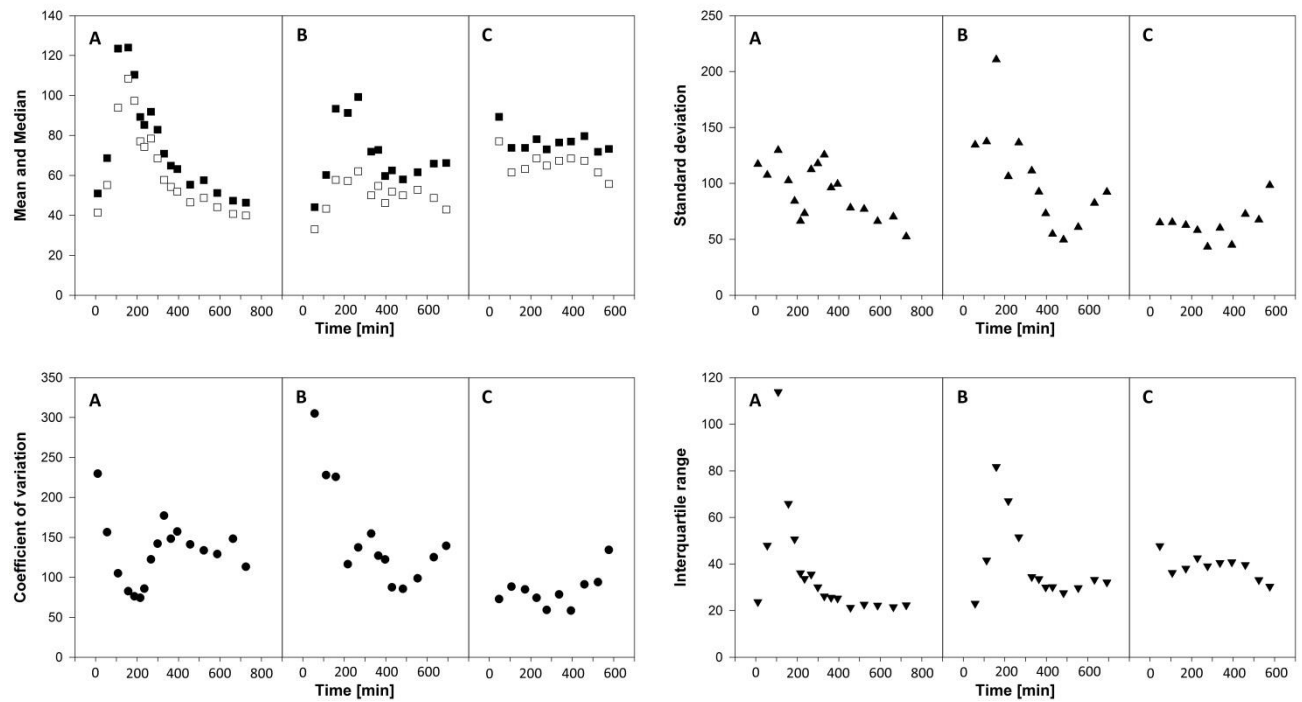


Figure 41 Box-plots showing the distribution of forward scatter per cell, for *B. subtilis* reporter strain AE099 cultivated in LB inoculated with a stationary phase LB pre-culture (A), in glucose chemically defined medium inoculated with a stationary phase LB pre-culture (B) or cultivated as a balanced cultures in glucose chemically defined medium (C).



**Figure 42** Descriptive statistics based on forward scatter per cell for *B. subtilis* reporter strain AE099 cultivated in LB inoculated with a stationary phase LB pre-culture (A), or in glucose chemically defined medium inoculated with a stationary phase LB pre-culture (B) or cultivated as a balanced cultures in glucose chemically defined medium (C). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 41.

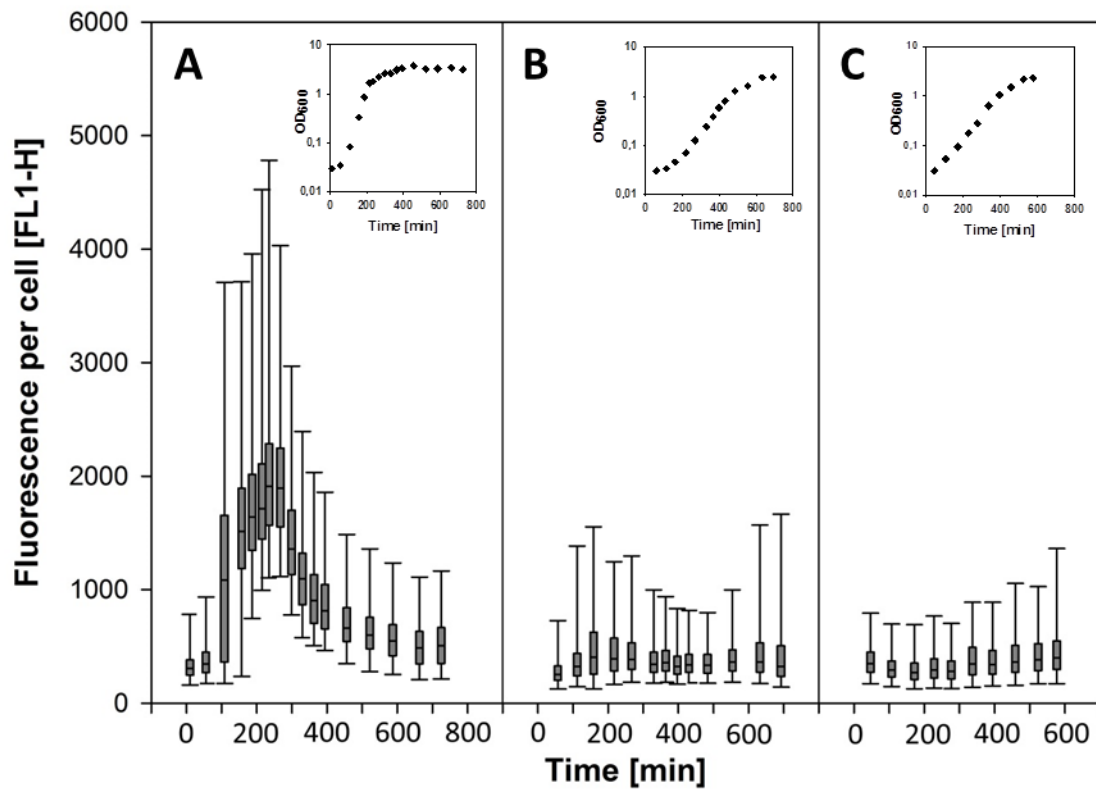
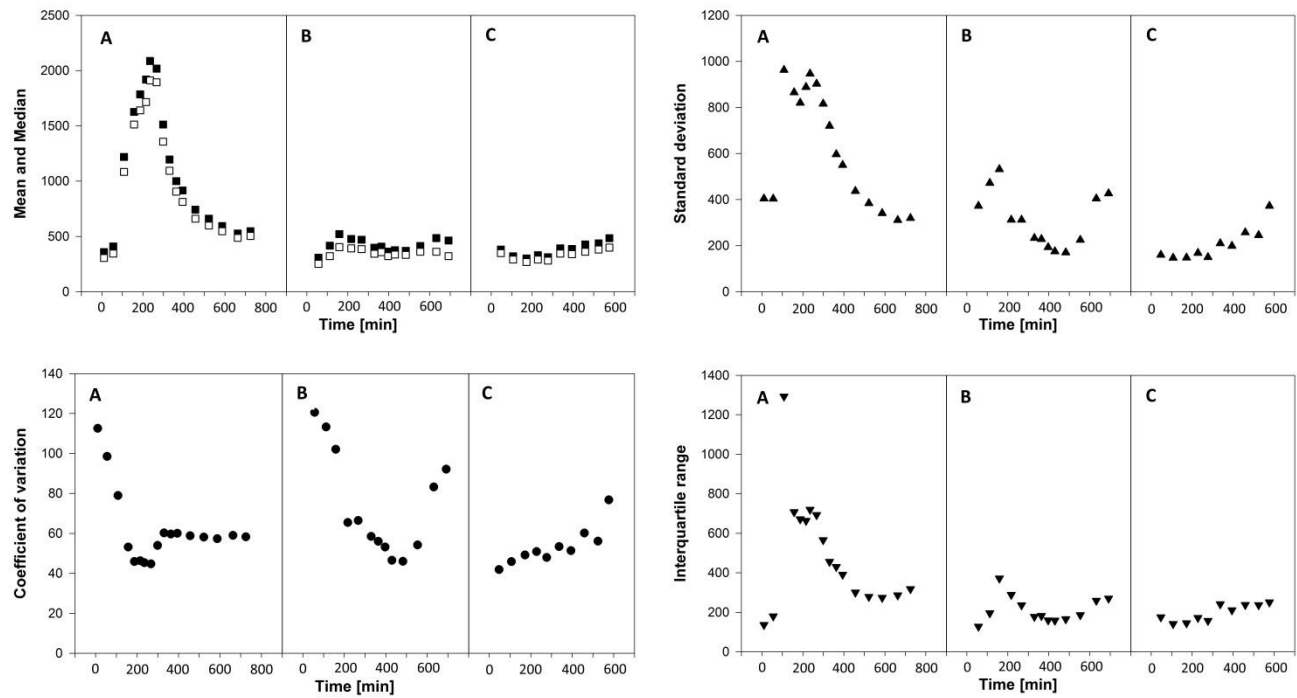


Figure 43 Box-plots showing the distribution of fluorescence per cell, for *B. subtilis* reporter strain AE099 cultivated in LB inoculated with a stationary phase LB pre-culture (A), in glucose chemically defined medium inoculated with a stationary phase LB pre-culture (B) or cultivated as a balanced cultures in glucose chemically defined medium (C).



**Figure 44** Descriptive statistics based on fluorescence per cell for *B. subtilis* reporter strain AE099 cultivated in LB inoculated with a stationary phase LB pre-culture (A), or in glucose chemically defined medium inoculated with a stationary phase LB pre-culture (B) or cultivated as a balanced cultures in glucose chemically defined medium (C). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 43.

### 11.2.3 *B. subtilis* cultivated in glucose chemically defined medium inoculated with either a stationary phase pre-culture or an exponentially growing pre-culture

This experiment, where a stationary phase pre-culture and an exponentially growing pre-culture of the *B. subtilis* reporter strain AE099 were used as inoculum, was made in duplicate. Specific cellular fluorescence computed from the flow cytometry data, can be found in the two manuscripts “heterogeneity at a single cell level in homogenous cultures” included in section 3 and “Statistical methods for assessment of heterogeneity in populations of isogenic bacteria analyzed by flow cytometry” included in section 11.1. In the sections below are the gated flow cytometry data (forward scatter per cell and fluorescence per shown) shown in box-plots and the descriptive statistics for both replicas are also included. The results shown in section 11.2.3.1 (replica A) can be linked to the manuscript included in section 3. The results shown in section 11.2.3.2 (replica B) can be linked to the manuscript in section 11.1

#### 11.2.3.1 Replica A

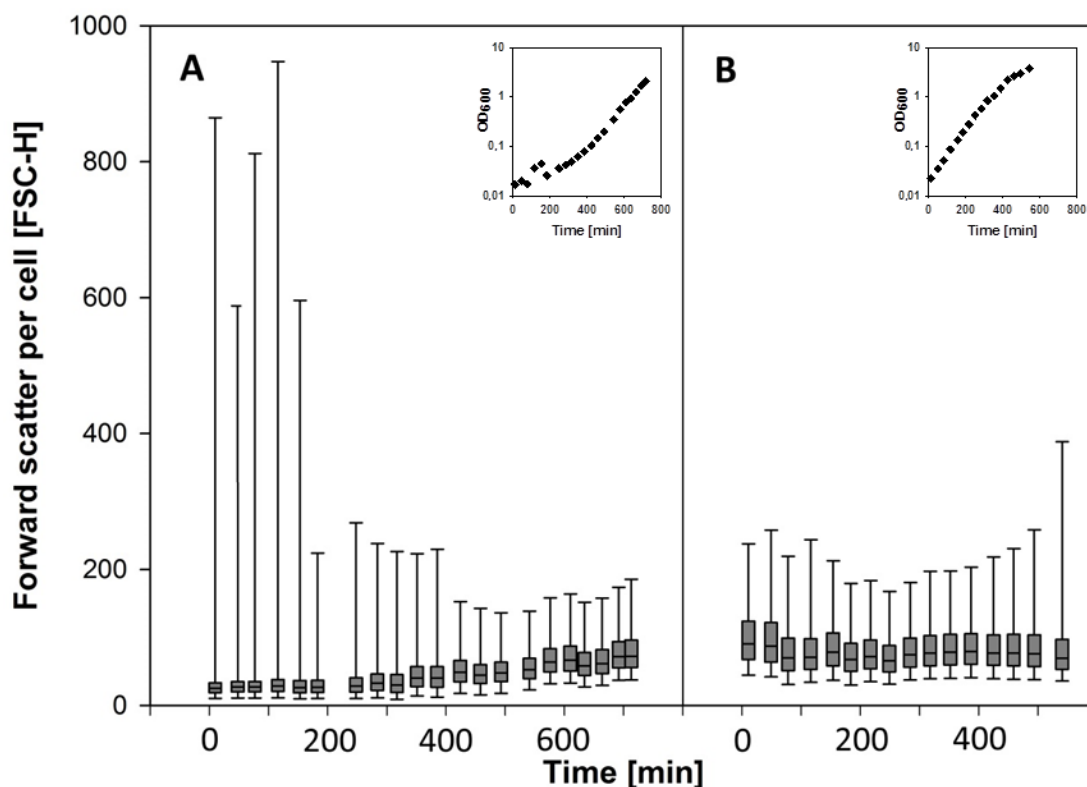
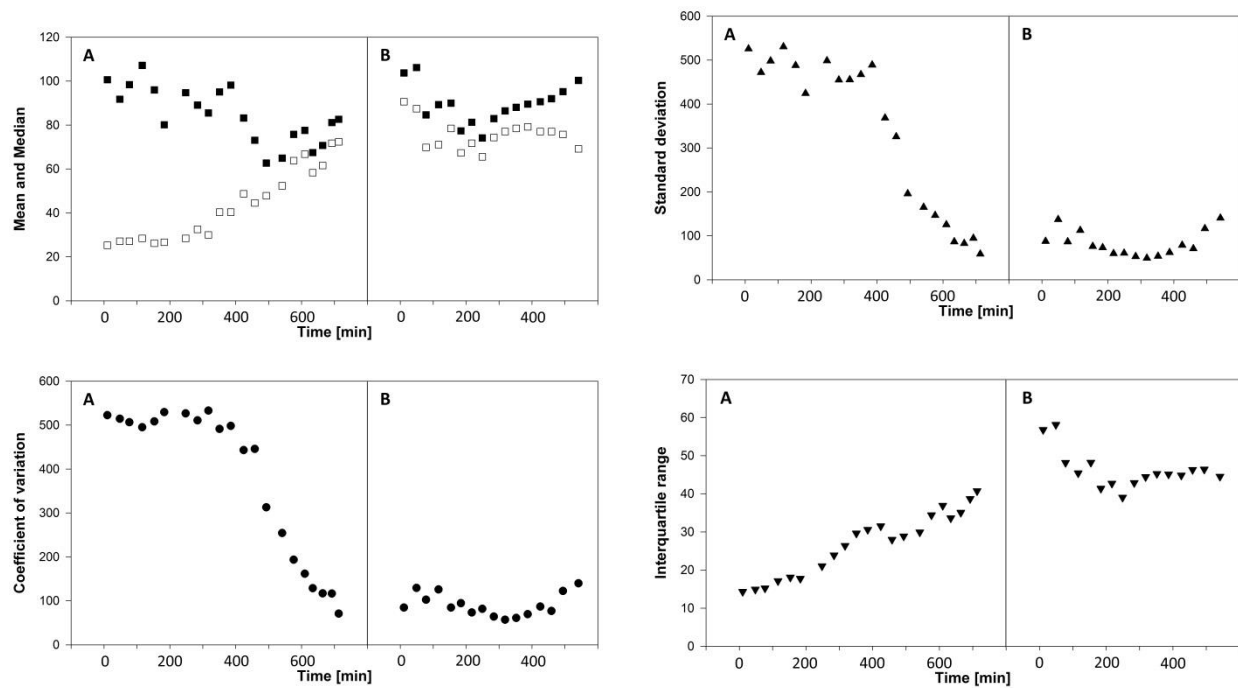


Figure 45 Box-plots showing the distribution of forward scatter per cell, for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose.



**Figure 46** Descriptive statistics based on forward scatter per cell for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 45.

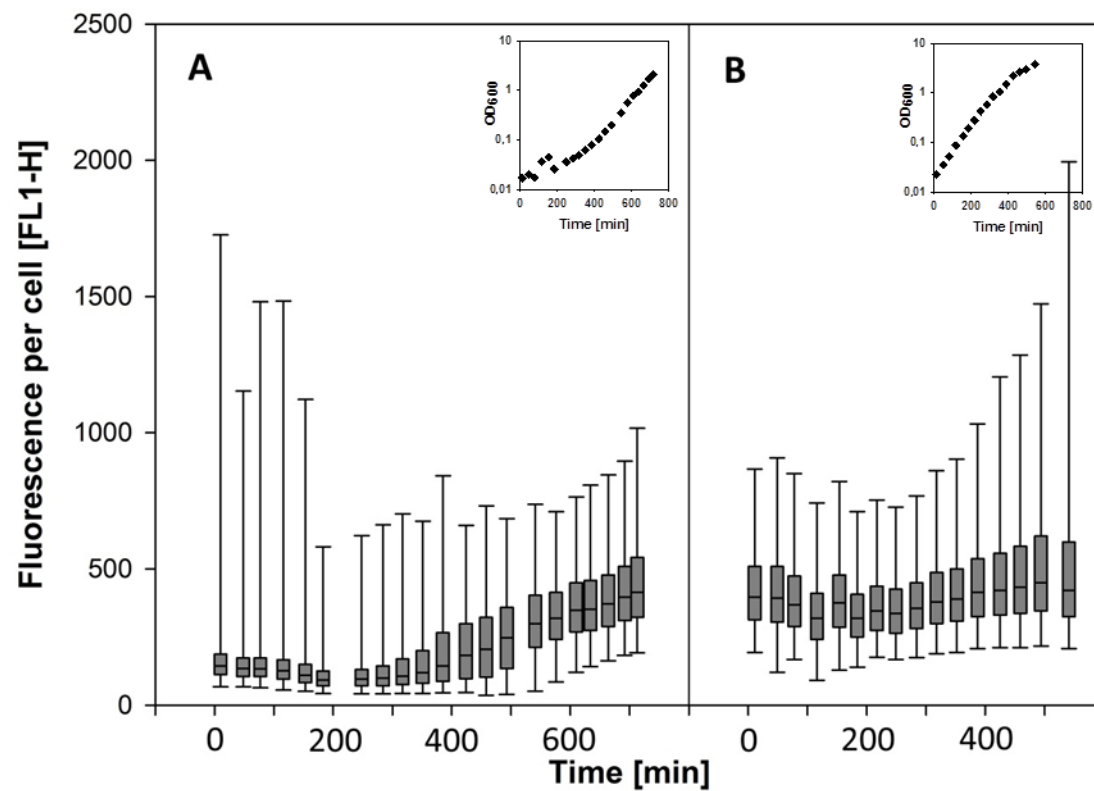
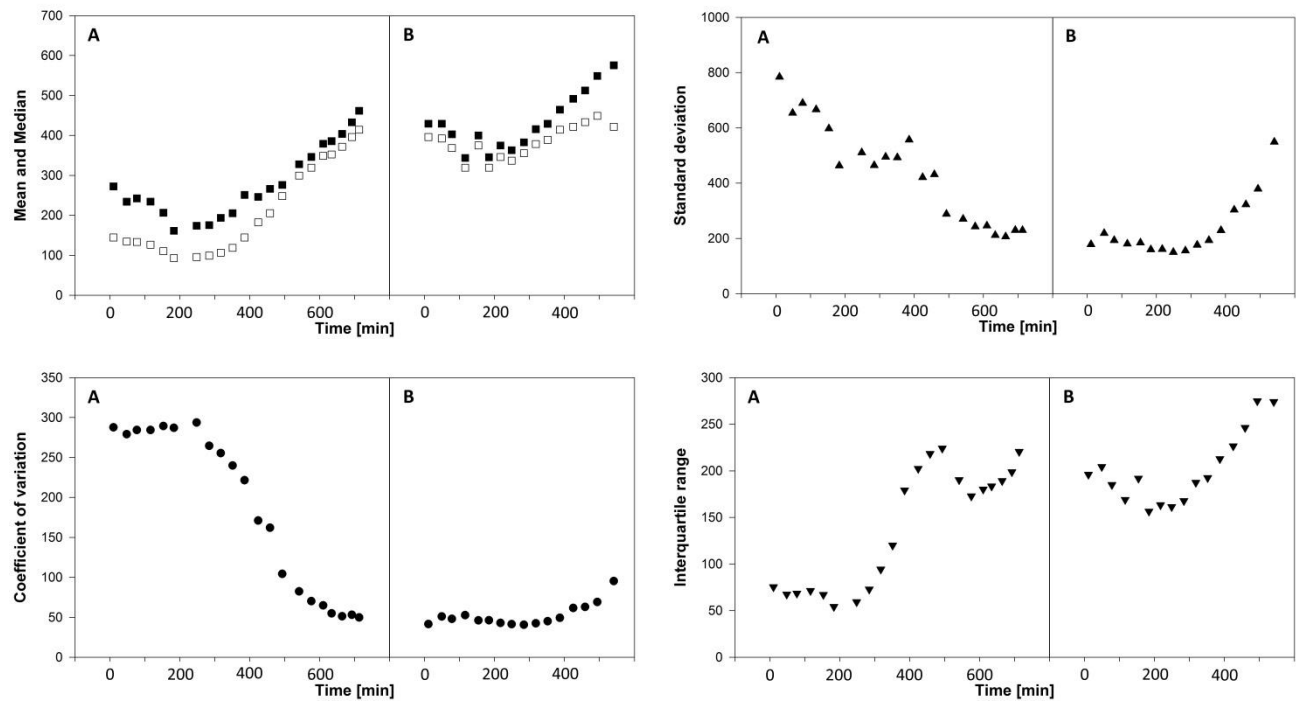


Figure 47 Box-plots showing the distribution of fluorescence per cell, for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose.



**Figure 48** Descriptive statistics based on fluorescence per cell for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 47.



### 11.2.3.2 Replica B

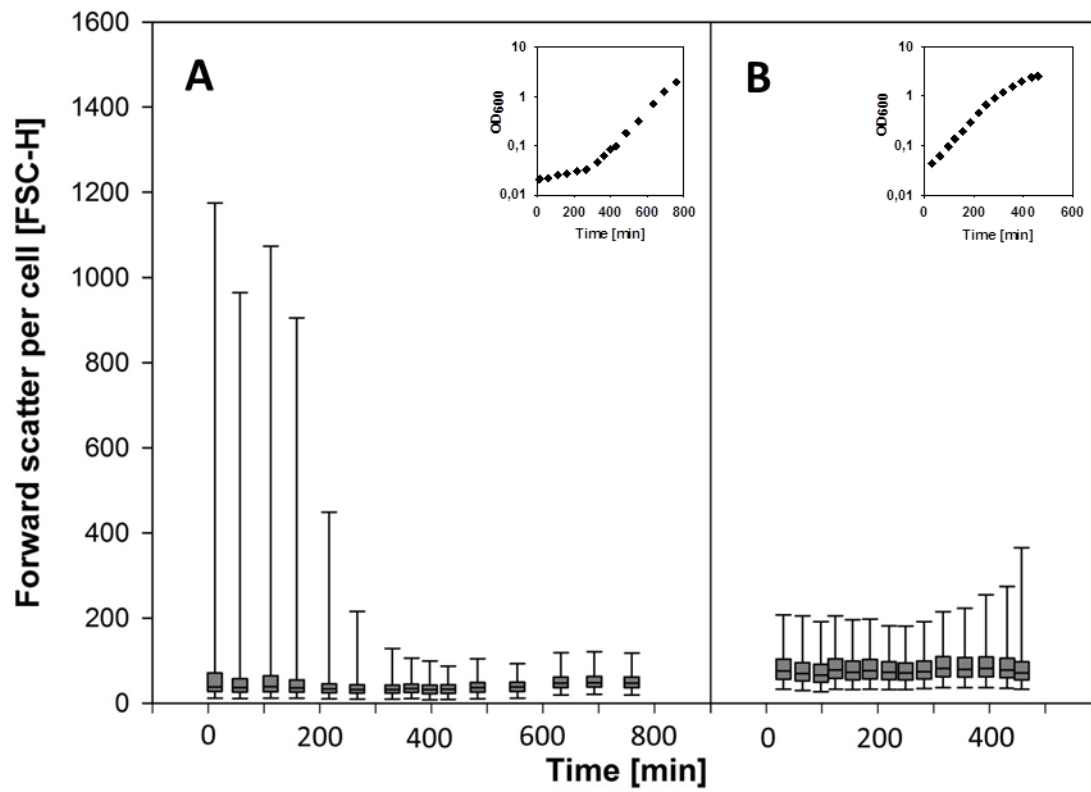
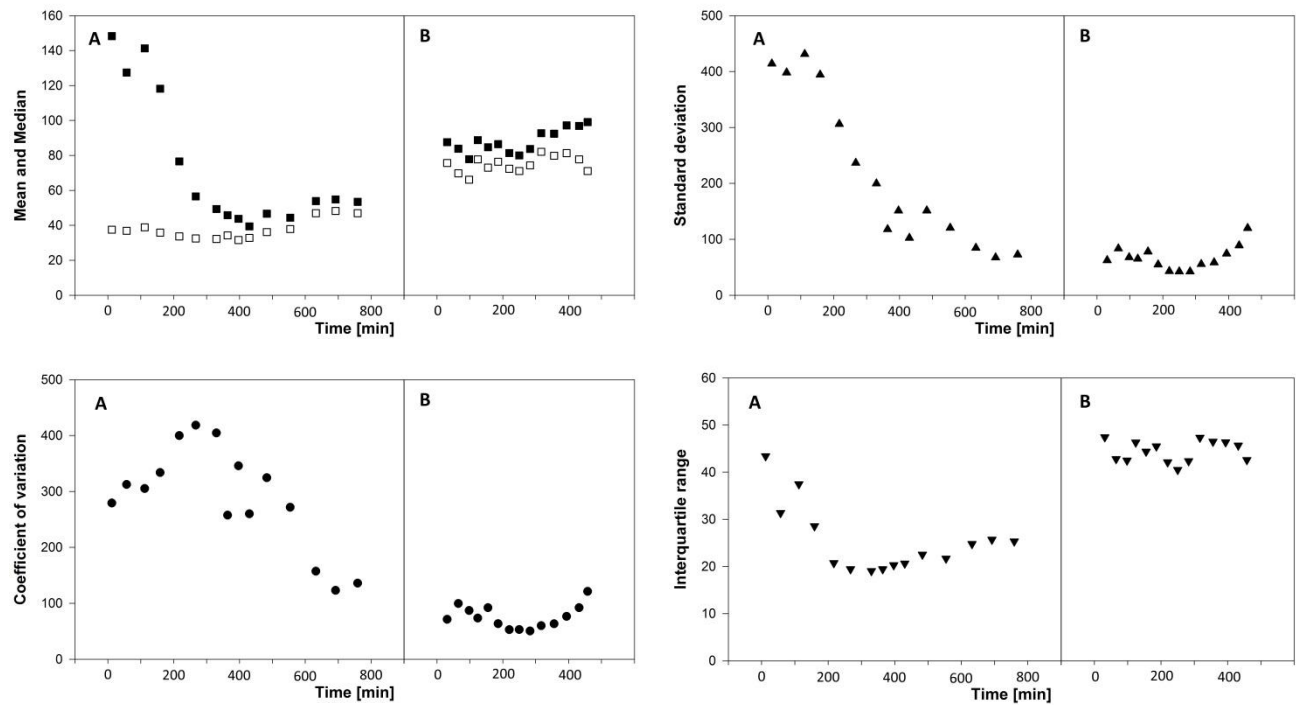


Figure 49 Box-plots showing the distribution of forward scatter per cell, for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose.



**Figure 50** Descriptive statistics based on forward scatter per cell for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 49.

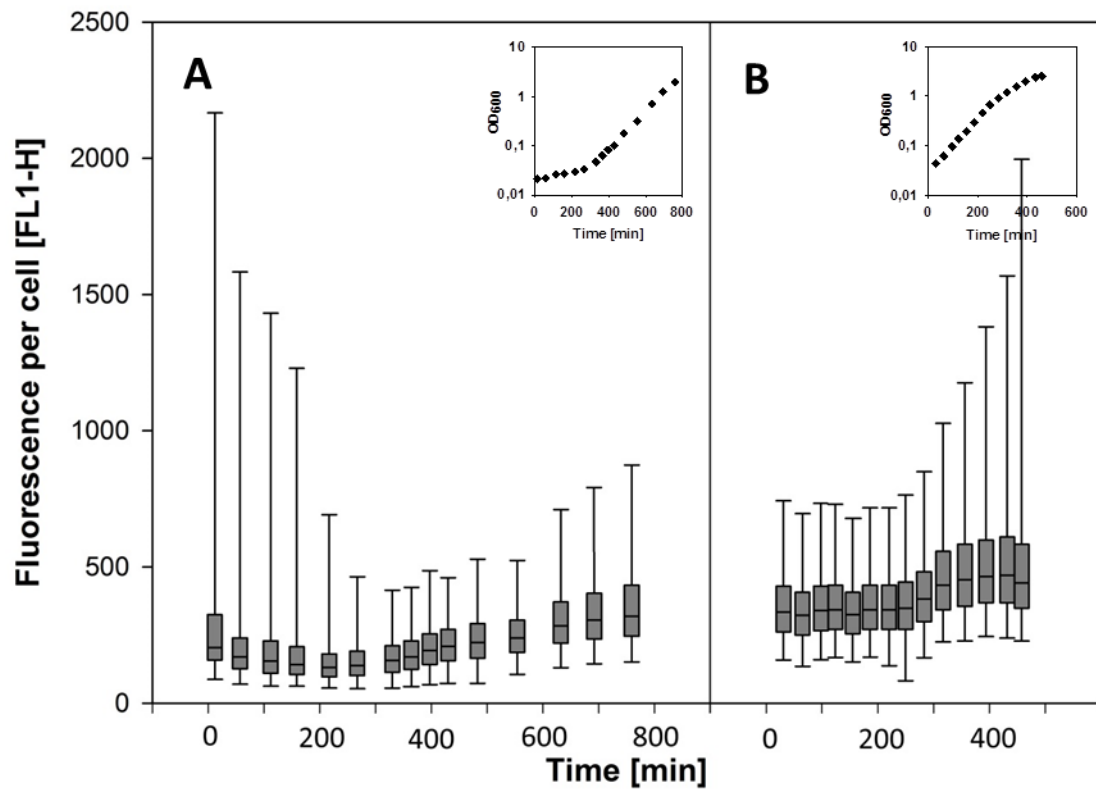
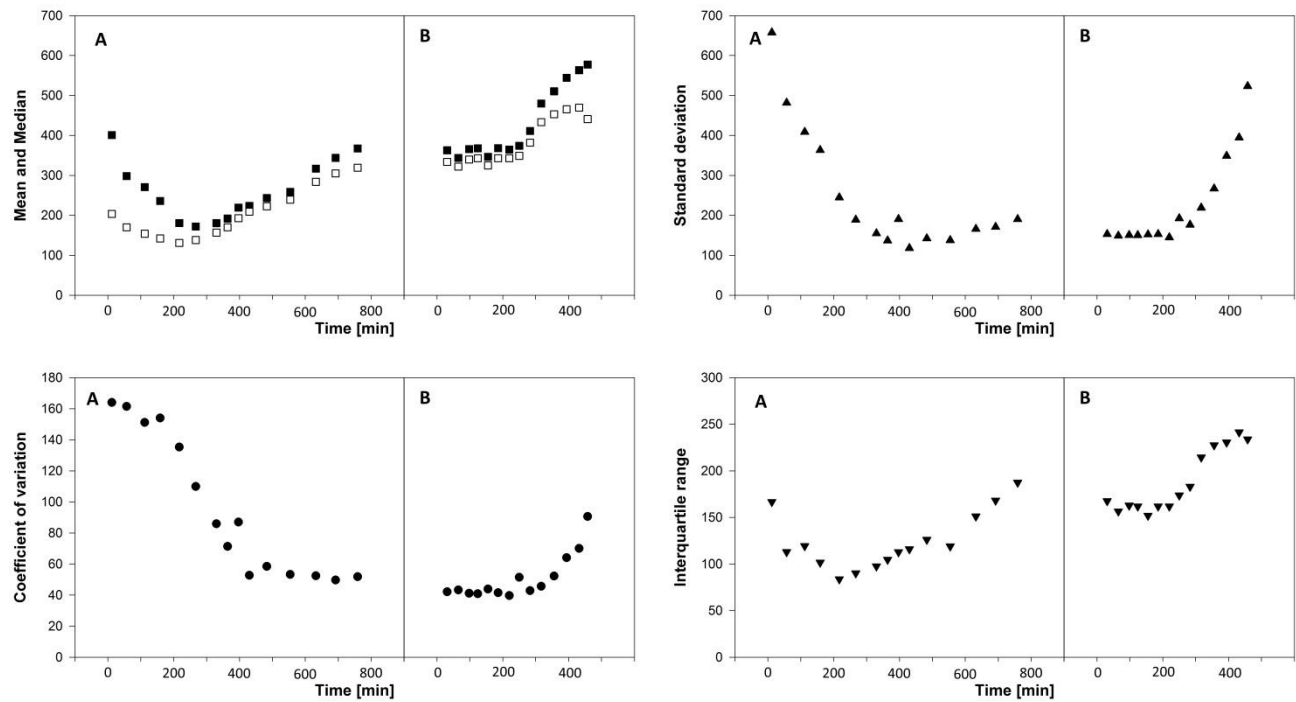


Figure 51 Box-plots showing the distribution of fluorescence per cell, for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose.

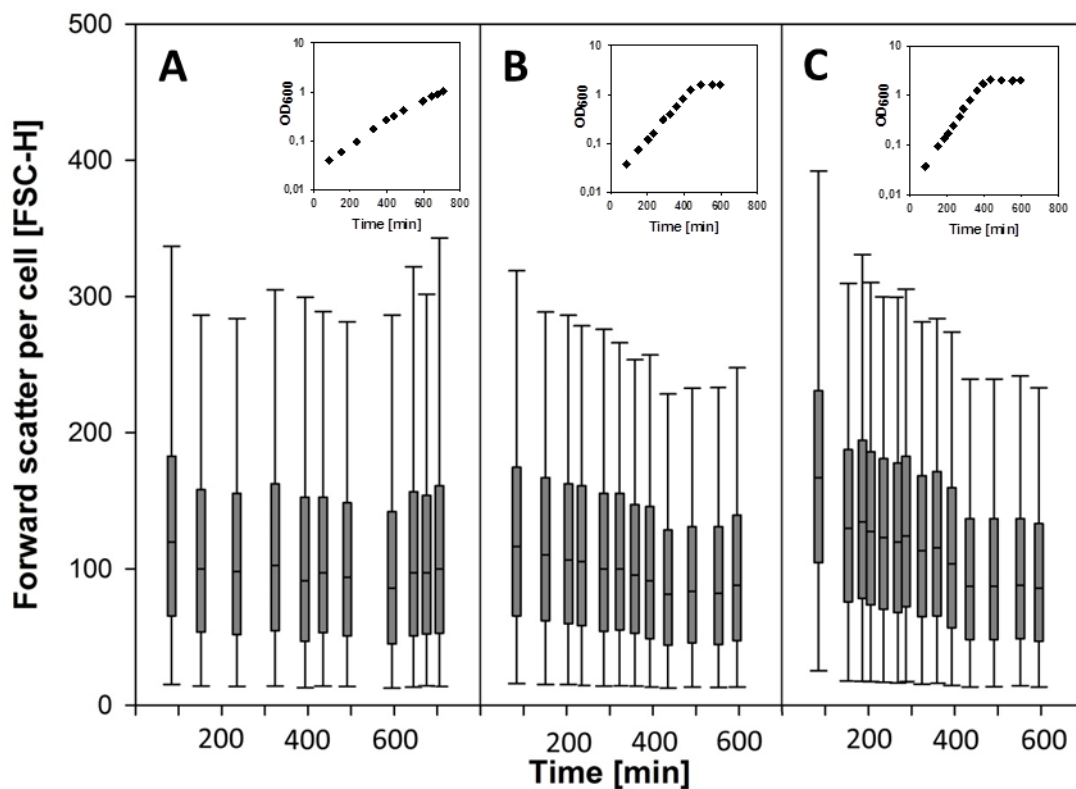


**Figure 52** Descriptive statistics based on fluorescence per cell for *B. subtilis* reporter strain AE099 cultivations inoculated with a stationary phase pre-culture (A), or a exponential growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium supplemented with glucose. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 51.

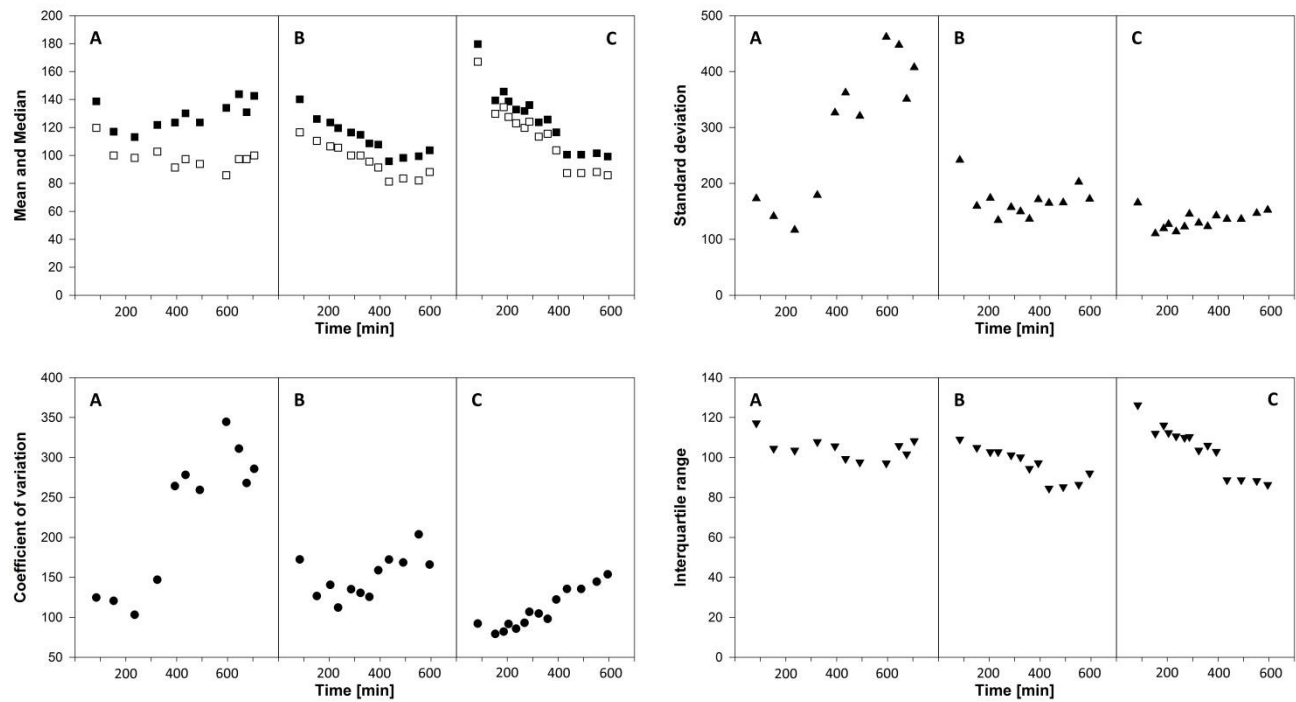
## 11.3 Supplementary materials – section 4

### 11.3.1 Heterogeneity determined in populations of *L. lactis* growing at three different rates

*L. lactis* GFP reporter strain AE072 was cultivated as balanced growing cultures in chemically defined medium supplemented with maltose, glucose or glucose plus nucleosides resulting in three different growth rates. Flow cytometry was used to analyze population heterogeneity. Box-plot showing the distribution for forward scatter per cell and fluorescence per cell is presented below, together with descriptive statistics calculated based on the flow cytometry data. Supplementary materials to section 0



**Figure 53** Box-plots showing the distribution of forward scatter per cell, determined for the *L. lactis* reporter strain AE072 when cultivated at three different growth rates. (A) AE072 cultivated in chemically defined medium supplemented with maltose, doubling time 133 minutes. (B) AE072 cultivated in chemically defined medium supplemented with glucose, doubling time 70 minutes. (C) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes.



**Figure 54** Descriptive statistics based on forward scatter per cell determined for the *L. lactis* reporter strain AE072 when cultivated at three different growth rates. (A) AE072 cultivated in chemically defined medium supplemented with maltose, doubling time 133 minutes. (B) AE072 cultivated in chemically defined medium supplemented with glucose, doubling time 70 minutes. (C) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 53.

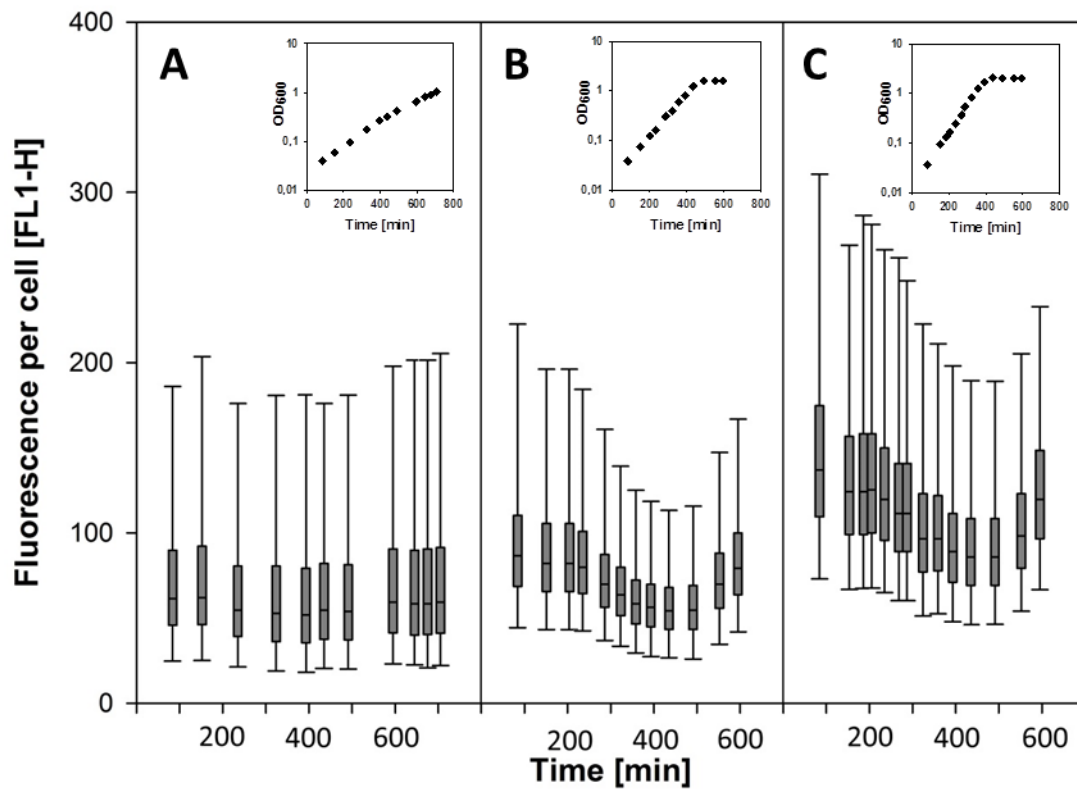
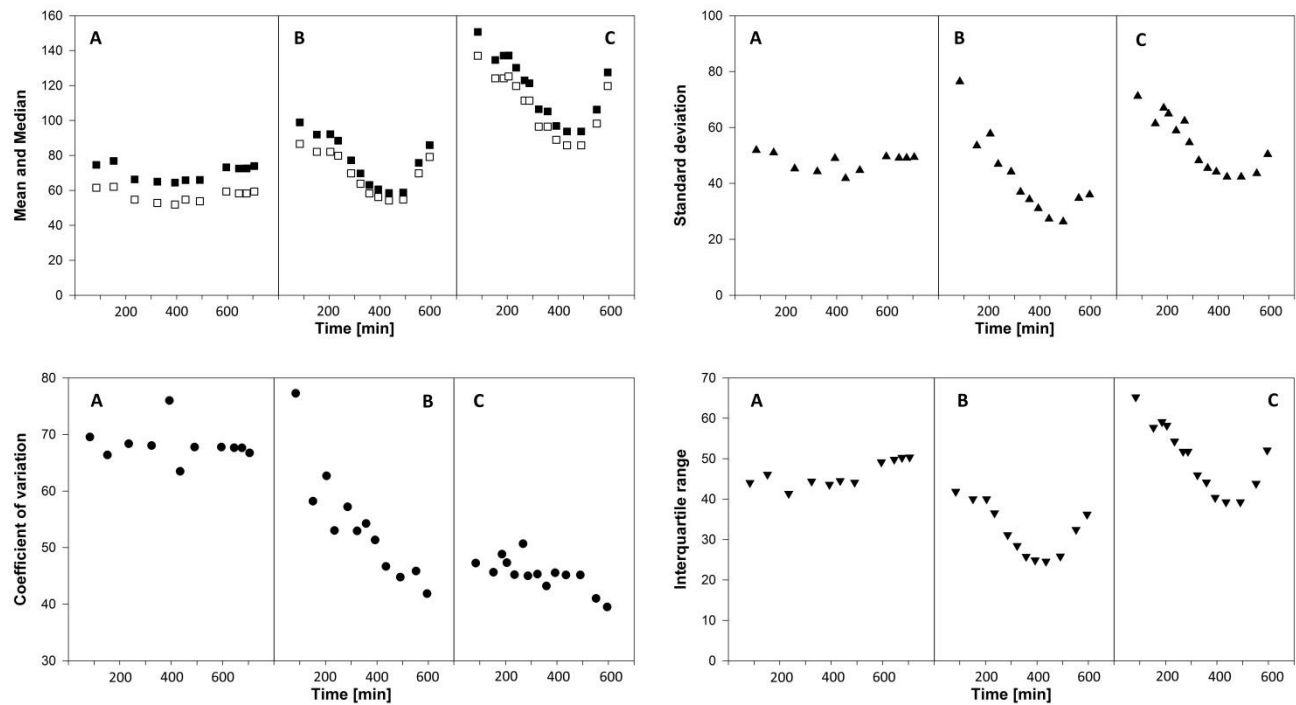


Figure 55 Box-plots showing the distribution of fluorescence per cell, determined for the *L. lactis* reporter strain AE072 when cultivated at three different growth rates. (A) AE072 cultivated in chemically defined medium supplemented with maltose, doubling time 133 minutes. (B) AE072 cultivated in chemically defined medium supplemented with glucose, doubling time 70 minutes. (C) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes.

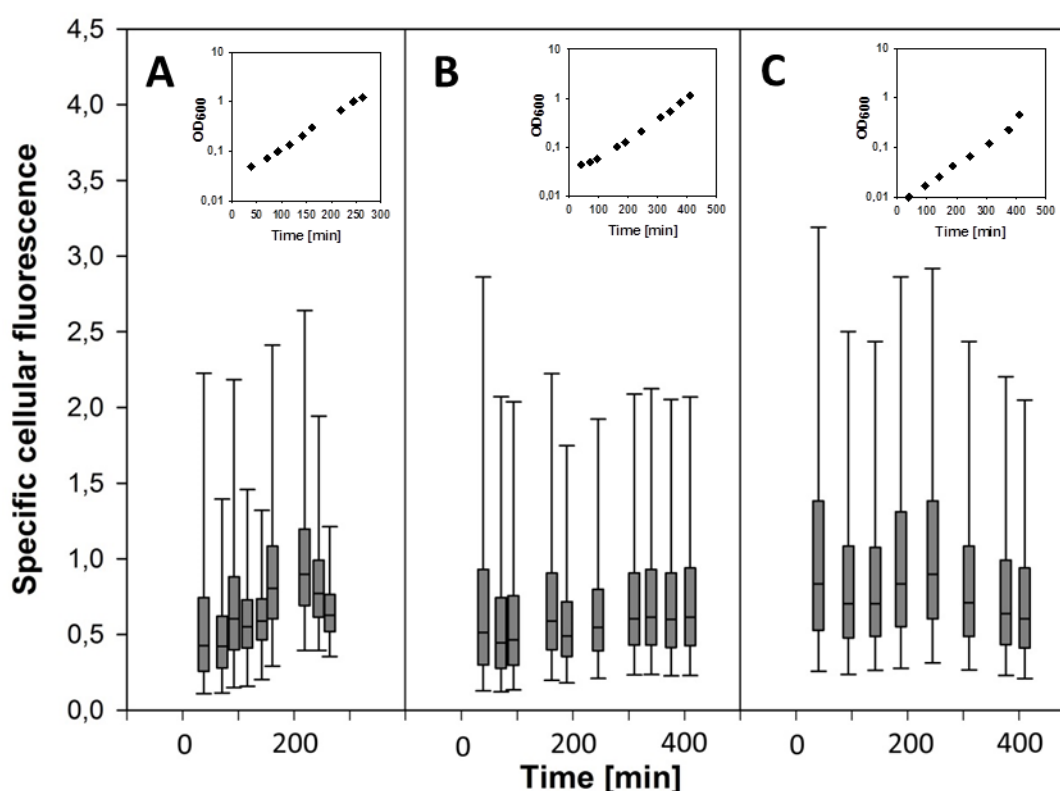


**Figure 56** Descriptive statistics based on fluorescence per cell determined for the *L. lactis* reporter strain AE072 when cultivated at three different growth rates. (A) AE072 cultivated in chemically defined medium supplemented with maltose, doubling time 133 minutes. (B) AE072 cultivated in chemically defined medium supplemented with glucose, doubling time 70 minutes. (C) AE072 cultivated in chemically defined medium supplemented with glucose and nucleosides, doubling time 54 minutes. (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 55

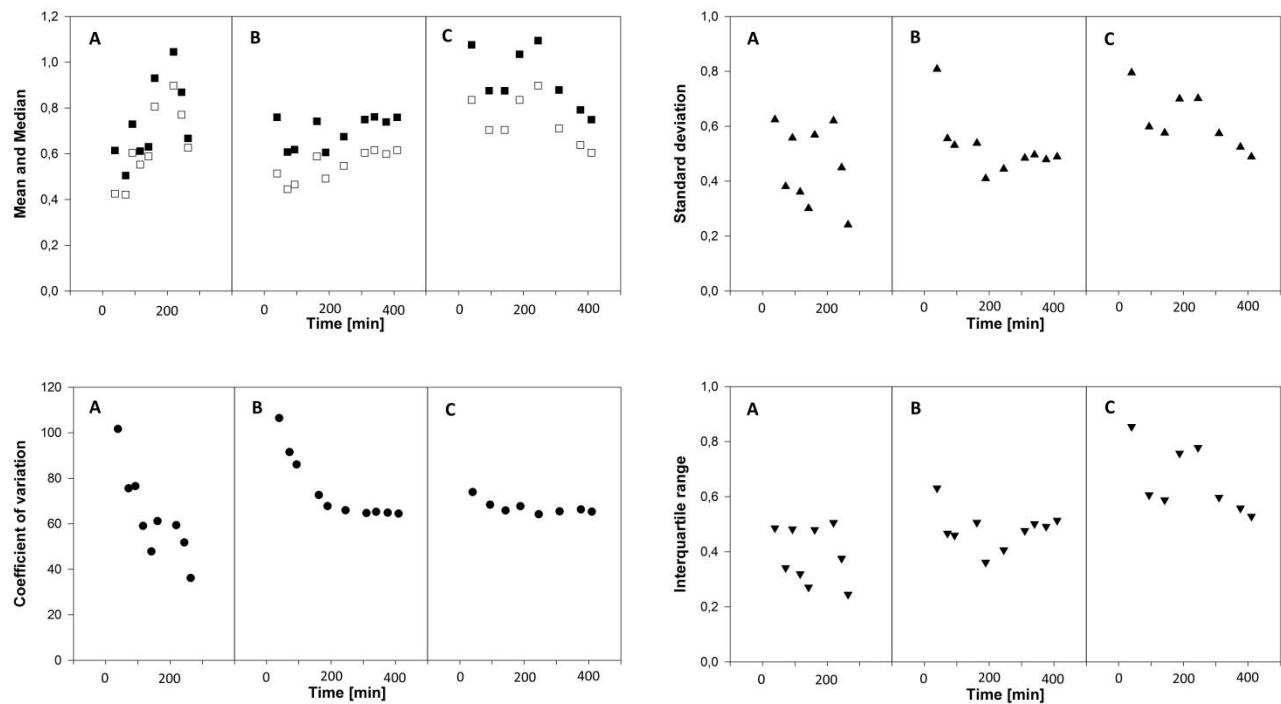


### 11.3.2 Heterogeneity determined in populations of *L. lactis* growing in complex or chemically defined medium

*L. lactis* GFP reporter strain AE072 was cultivated in complex (GM17) medium and chemically defined medium (SALM) inoculated with a stationary phase complex medium pre-culture. For comparison was a balanced growing chemically defined medium culture included. Flow cytometry was used to analyze population heterogeneity. Box-plot showing the distribution for forward scatter per cell and fluorescence per cell is presented below, together with descriptive statistics calculated based on the flow cytometry data. Supplementary materials to section 4.3.3



**Figure 57** Box-plots showing the distribution of specific cellular fluorescence observed for the *L. lactis* reporters strain AE072. (A) Complex medium culture inoculated with a stationary phase complex medium pre-culture. (B) Chemically defined medium (SALM supplemented with glucose) culture inoculated with a stationary phase complex medium pre-culture. (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose).



**Figure 58** Descriptive statistics based on specific cellular fluorescence observed for the *L. lactis* reporter strain AE072. (A) Complex medium culture inoculated with a stationary phase complex medium pre-culture. (B) Chemically defined medium (SALM supplemented with glucose) culture inoculated with a stationary phase complex medium pre-culture. (C) A balanced growing culture in chemically defined medium (SALM supplemented with glucose). The associated growth curves are in Figure 57. (Top left) Mean: black squares, median white squares, (top right) standard deviation: triangle up, (bottom left) coefficient of variation: black circle, (bottom right) interquartile range: triangle down.

## 11.4 Supplementary material – section 5

The *S. cerevisiae* GFP reporter strain FE440 was cultivated in chemically defined medium inoculated with either a stationary phase pre-culture or an exponentially growing pre-culture. The experiment was repeated once. The distributions of the computed parameter specific cellular fluorescence measured for replica A is shown in section 5.3.2 as box-plot. The descriptive statistics for replica A specific cellular fluorescence are also included in this section. Below are box-plots and descriptive statistics for the flow cytometry measurements of the forward scatter per cell and fluorescence per cell included, as well as the data for replica B specific cellular fluorescence.

### 11.4.1.1 Replica A

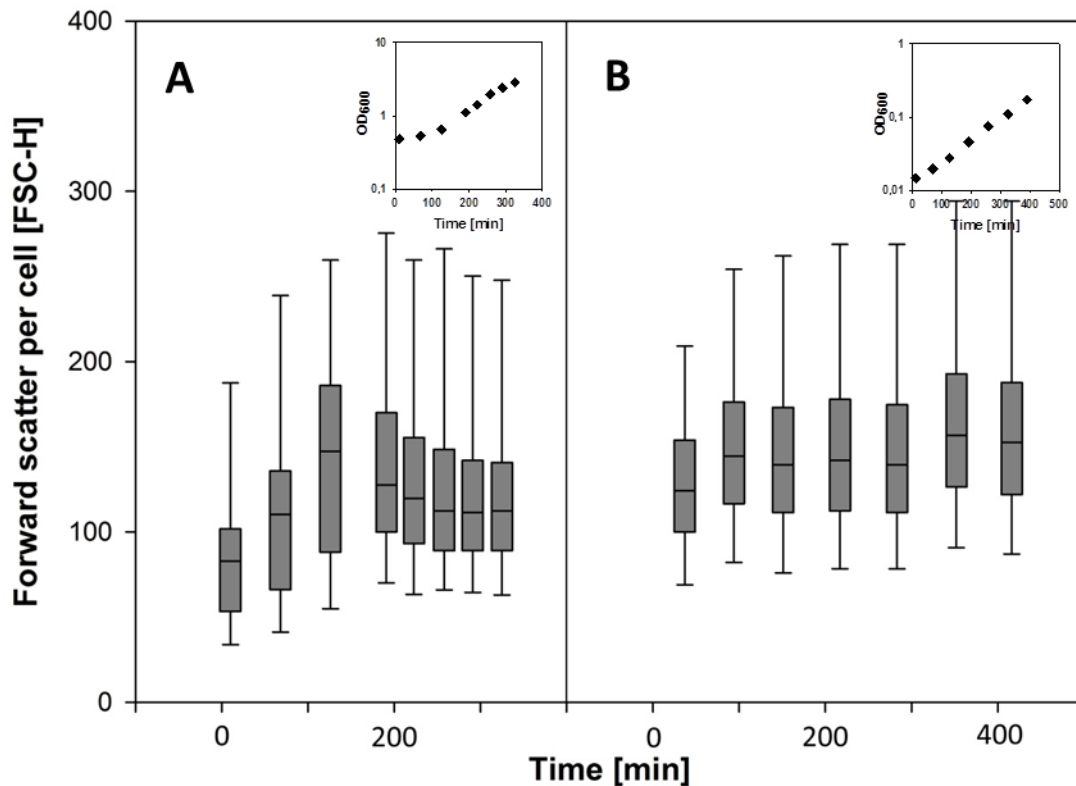
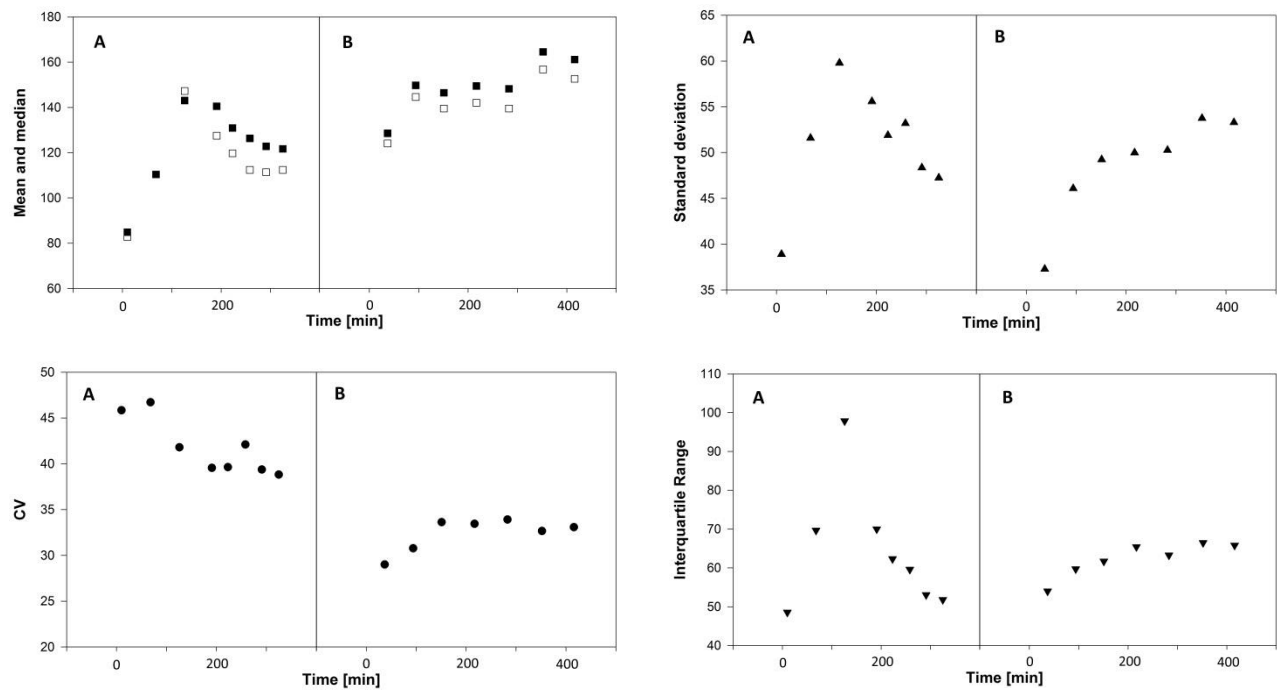


Figure 59 Box-plots showing the distribution of forward scatter per cell measured in cultivations of *S. cerevisiae* reporter strain 440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). Growth rates in exponential phase: (A)  $0.27 \text{ h}^{-1}$  ( $R^2=0.99$ ) and (B)  $0.39 \text{ h}^{-1}$  ( $R^2=0.99$ ).



**Figure 60** Descriptive statistics based on forward scatter per cell measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 59.

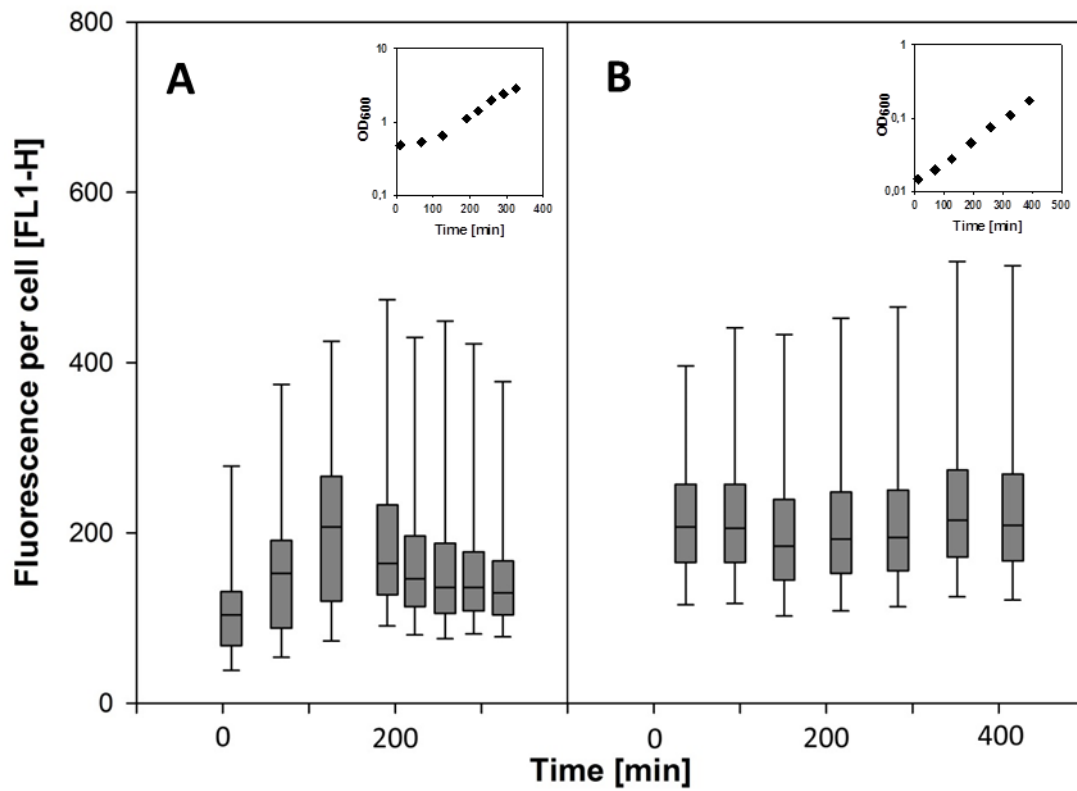
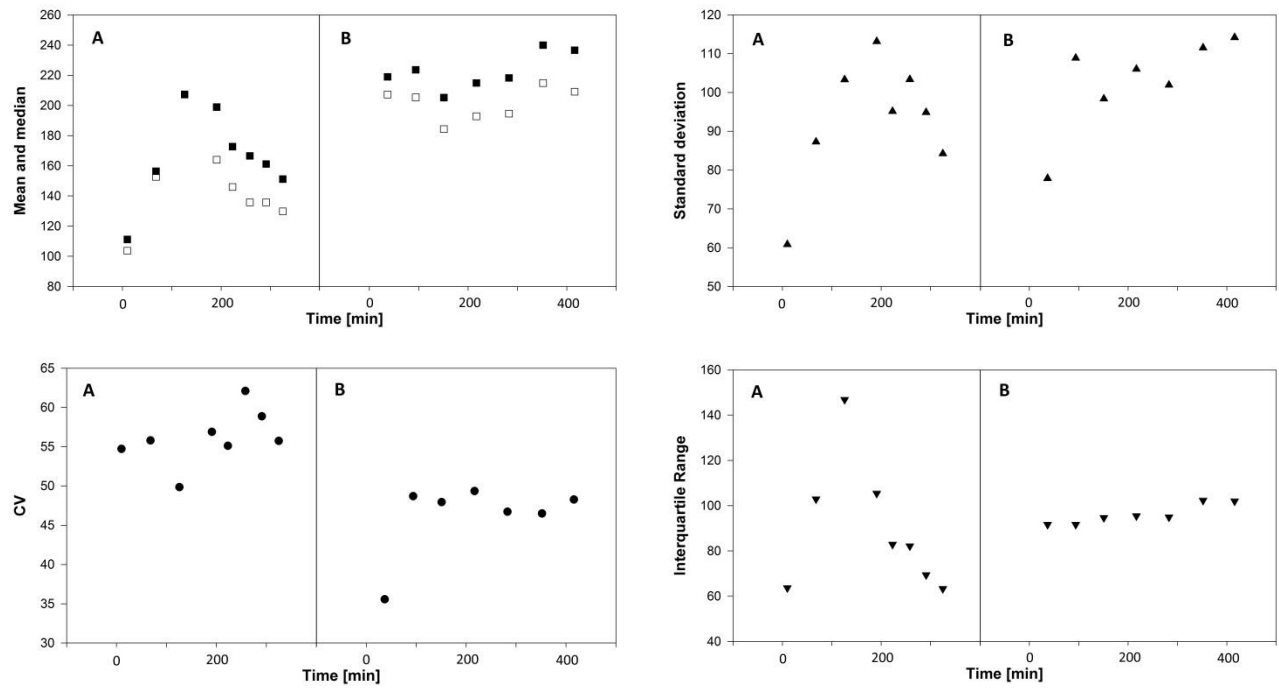
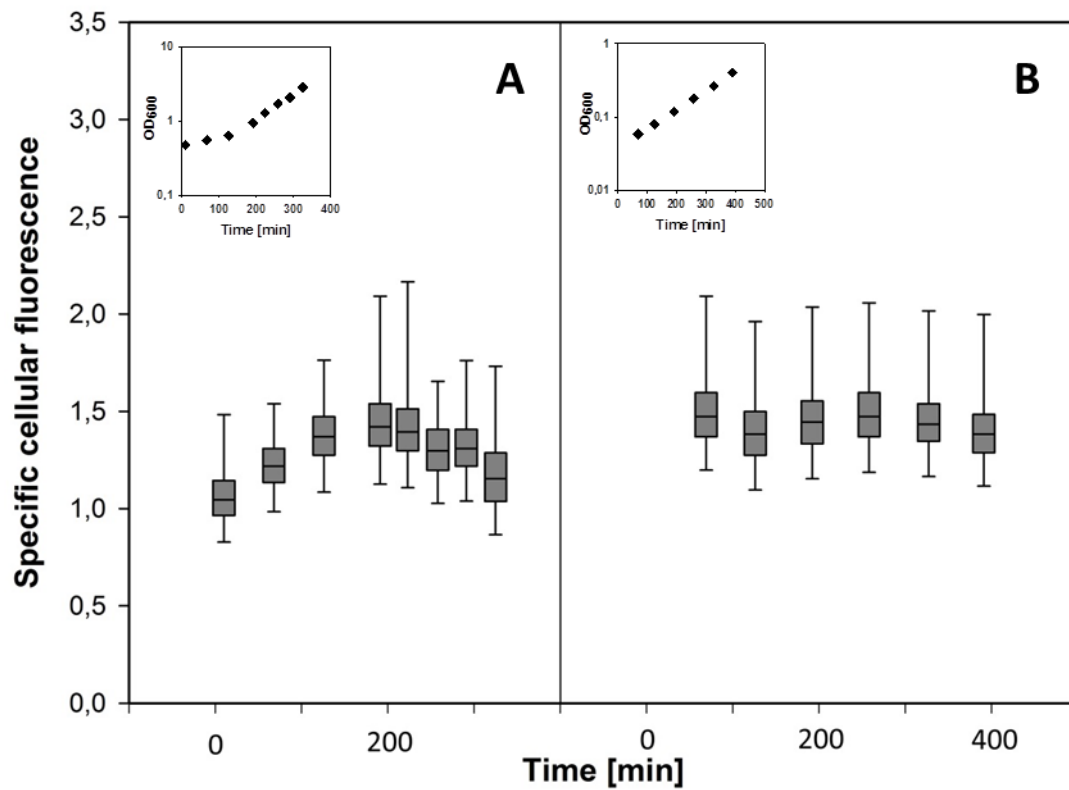


Figure 61 Box-plots showing the distribution of fluorescence per cell measured in cultivations of *S. cerevisiae* reporter strain 440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). Growth rates in exponential phase: (A) 0.27 h<sup>-1</sup> (R<sup>2</sup>=0.99) and (B) 0.39 h<sup>-1</sup> (R<sup>2</sup>=0.99).

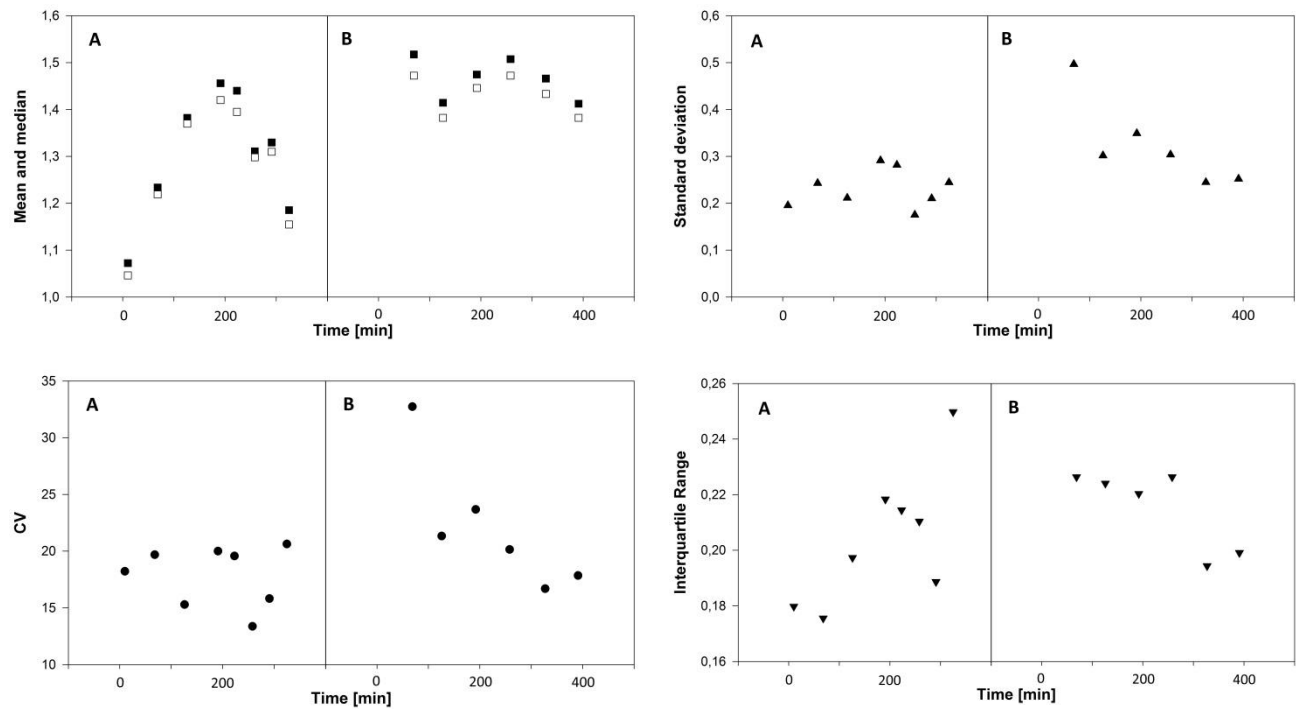


**Figure 62** Descriptive statistics based on fluorescence per cell measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 61.

#### 11.4.1.2 Replica B



**Figure 63** Box-plot showing the distribution of specific cellular fluorescence (single cell growth rate) during cultivation with the yeast reporter strain FE440 cultivated in chemically defined medium inoculated with either a stationary phase overnight culture (A) or an exponentially growing culture (B). Growth rates in exponential phase: (A)  $0.29 \text{ h}^{-1}$  ( $R^2=0.98$ ) and (B)  $0.35 \text{ h}^{-1}$  ( $R^2=0.99$ ).



**Figure 64** Descriptive statistics based on specific cellular fluorescence measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. Growth curves can be seen in Figure 63.



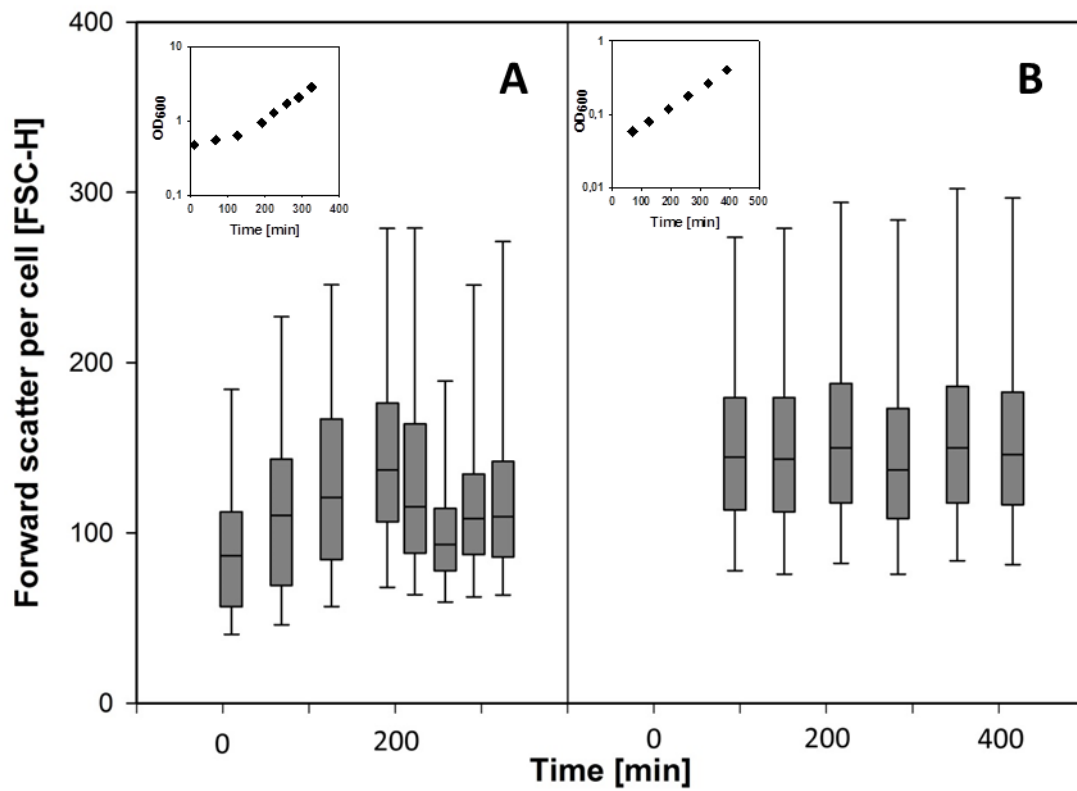
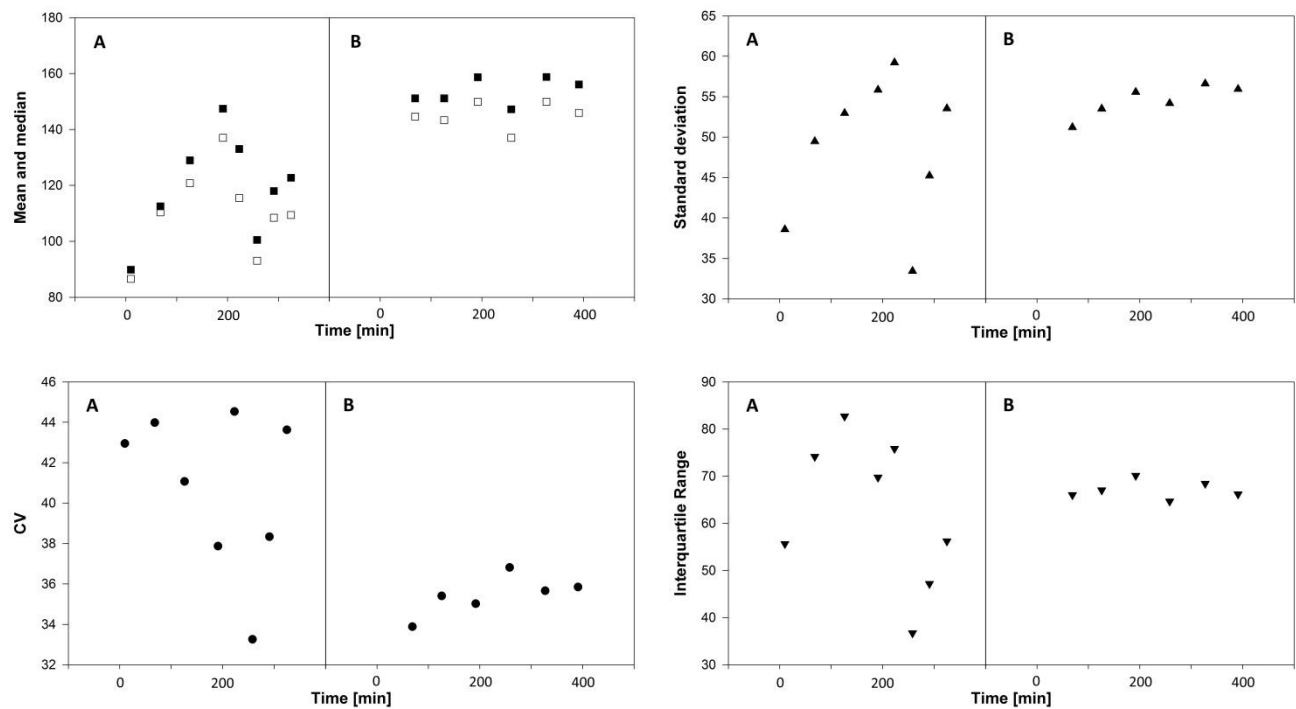


Figure 65 Box-plots showing the distribution of forward scatter per cell measured in cultivations of *S. cerevisiae* reporter strain 440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). Growth rates in exponential phase: (A)  $0.29 \text{ h}^{-1}$  ( $R^2=0.98$ ) and (B)  $0.35 \text{ h}^{-1}$  ( $R^2=0.99$ ).



**Figure 66** Descriptive statistics based on forward scatter per cell measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. The associated growth curves are in Figure 65.

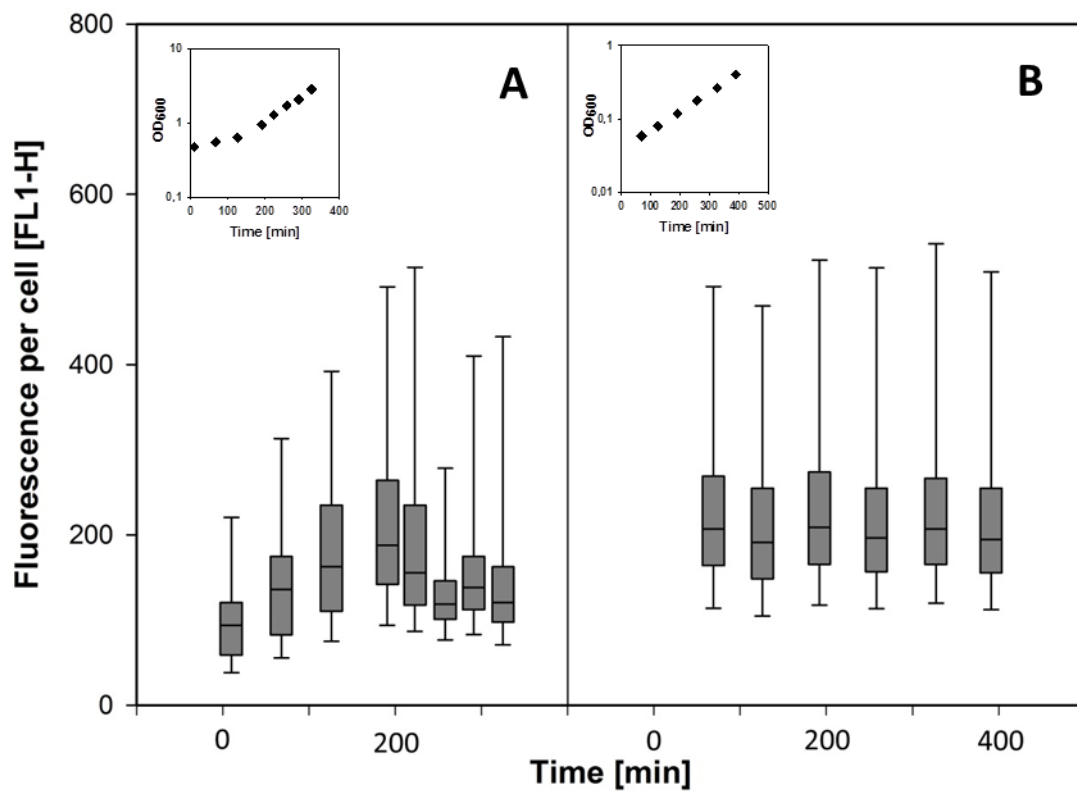
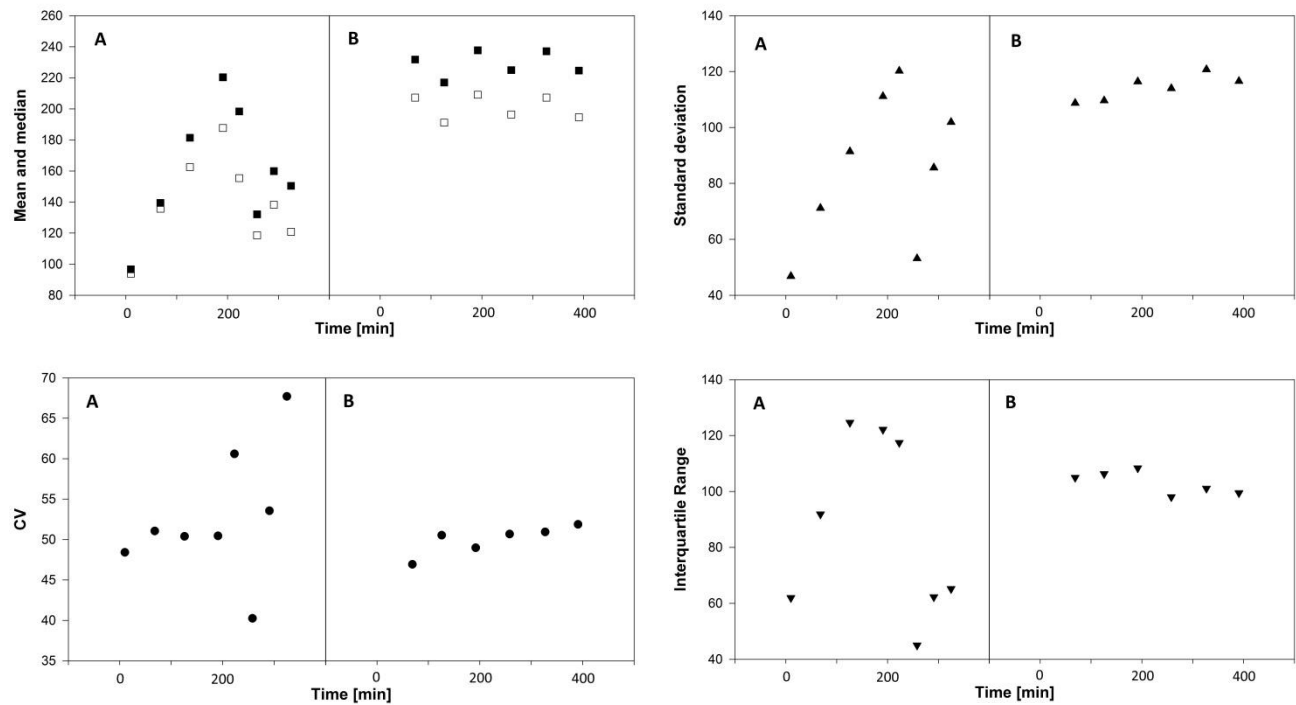


Figure 67 Box-plots showing the distribution of fluorescence per cell measured in cultivations of *S. cerevisiae* reporter strain 440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). Growth rates in exponential phase: (A)  $0.29 \text{ h}^{-1}$  ( $R^2=0.98$ ) and (B)  $0.35 \text{ h}^{-1}$  ( $R^2=0.99$ ).



**Figure 68** Descriptive statistics based on fluorescence per cell measured in cultures of the *S. cerevisiae* reporter strain FE440 inoculated with a stationary phase pre-culture (A) or an exponentially growing pre-culture (B). Both cultures and pre-cultures were propagated in chemically defined medium (CBS medium). (Top left) Mean: black squares, median white squares. (Top right) standard deviation: triangle up. (Bottom left) coefficient of variation: black circle. (Bottom right) interquartile range: triangle down. The associated growth curves are in Figure 67.

## 11.5 Supplementary material – section 6

*B. subtilis* was propagated as a balanced growing culture at 37°C, the incubation temperature was abrupt increased to 48°C exposing the culture to heat stress. The results of the analysis of this culture is shown below, together with the results determined for a balanced growing culture propagated at 37°C, included as a reference. The distribution determined for forward scatter per cell and fluorescence per cell is shown in box-plots below. Descriptive statistics describing the variation in the data is included in Figure 70 and Figure 72.

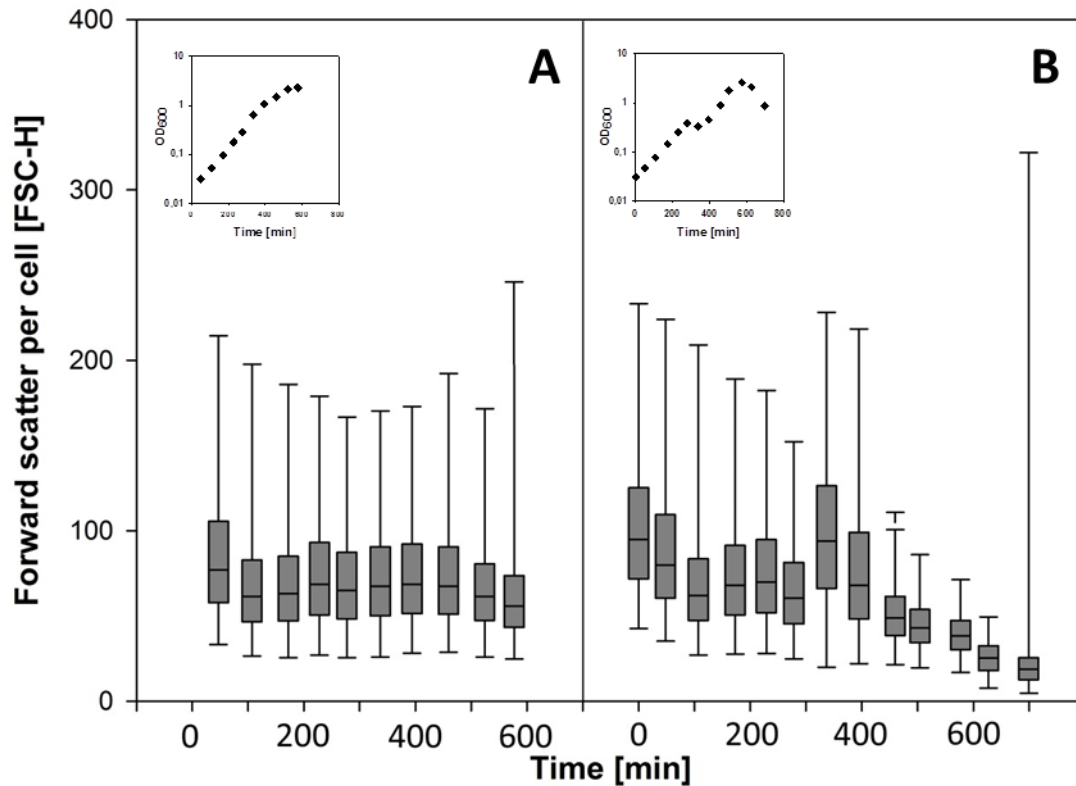
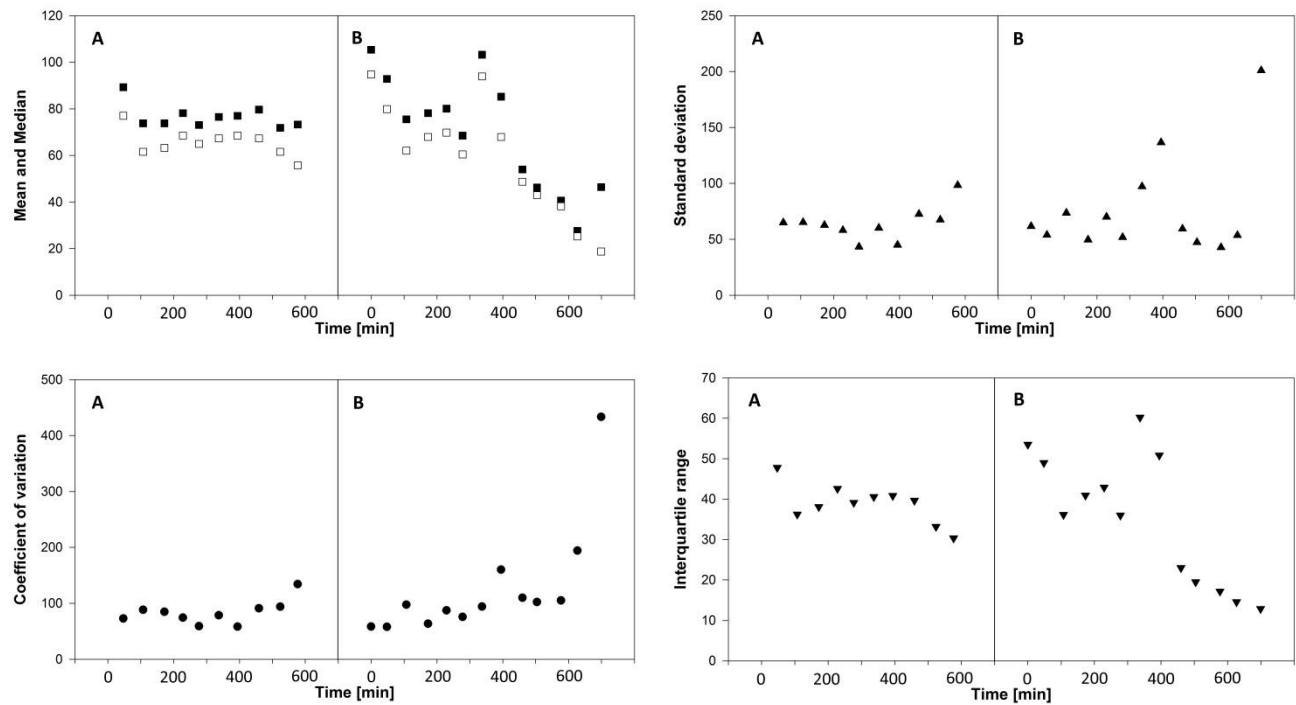


Figure 69 Box-plots showing the distribution of forward scatter per cell measured for two cultures of *B. subtilis* strain AE099. (A) A balanced growing culture propagated at 37°C was included as a control. (B) A balanced growing culture propagated at 37°C was, 278 minutes after inoculation, exposed to thermal stress by increasing the incubation temperature to 48°C.



**Figure 70** Descriptive statistics based on forward scatter per cell measured for the *B. subtilis* reporter strain AE099. (A) Balanced growing culture propagated at 37°C. (B) Balanced growing culture propagated at 37°C and 278 minutes after inoculation exposed to thermal stress at 48°C (Top left) mean: black squares and median: white squares. (Top right) standard deviation: triangles. (Bottom left) coefficient of variation: diamonds. (Bottom right) interquartile range: triangle, down. The descriptive statistics can be linked to the growth curves presented in Figure 69.

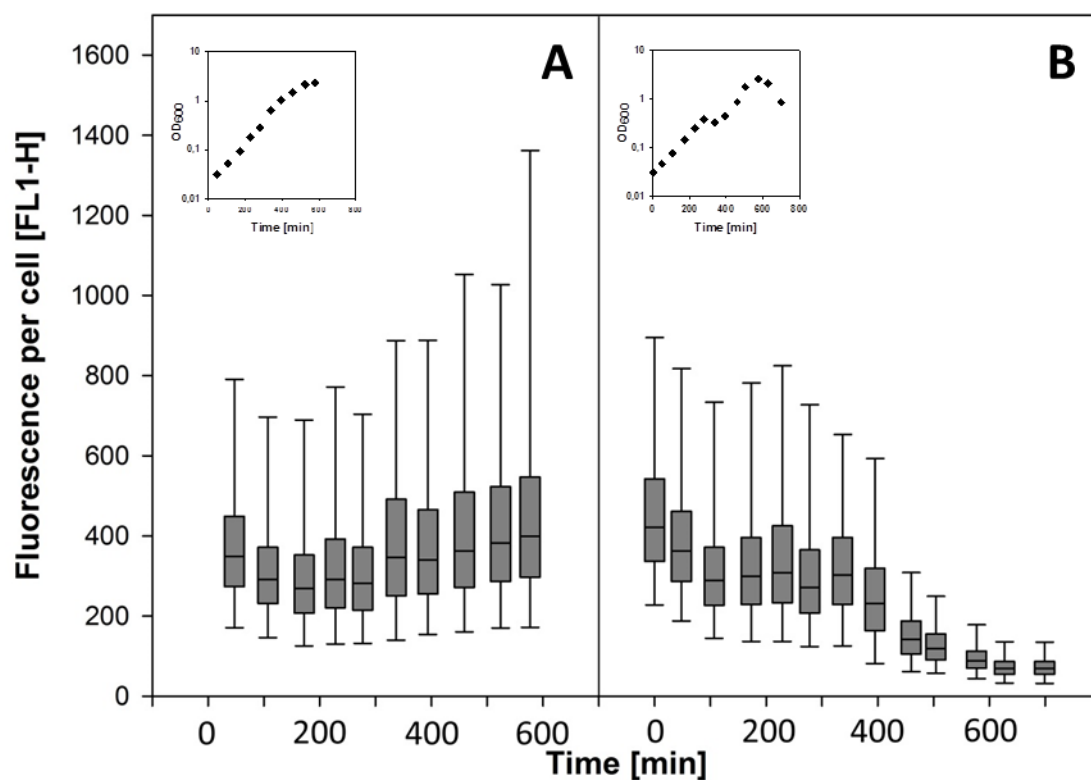
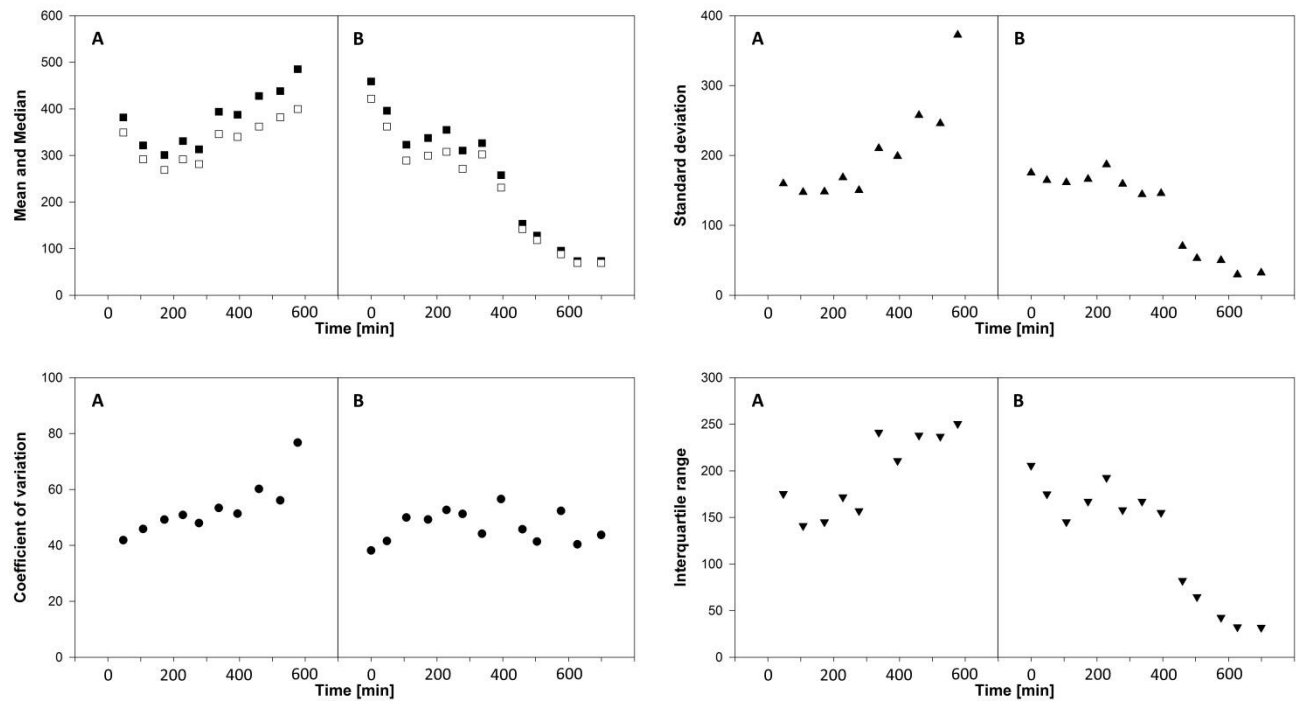


Figure 71 Box-plots showing the distribution of fluorescence per cell measured for two cultures of *B. subtilis* strain AE099. (A) A balanced growing culture propagated at 37°C was included as a control. (B) A balanced growing culture propagated at 37°C was, 278 minutes after inoculation, exposed to thermal stress by increasing the incubation temperature to 48°C.



**Figure 72** Descriptive statistics based on fluorescence per cell measured for the *B. subtilis* reporter strain. (A) Balanced growing culture propagated at 37°C. (B) Balanced growing culture propagated at 37°C, 278 minutes after inoculation exposed to thermal stress at 48°C (Top left) mean: black squares and median: white squares. (Top right) standard deviation: triangles. (Bottom left) coefficient of variation: diamonds. (Bottom right) interquartile range: triangle, down. The descriptive statistics can be linked to the growth curves presented in Figure 71.